

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 13:16:38 ; Search time 35 Seconds
(without alignments)
87.565 Million cell updates/sec

Title: US-09-641-034-47

Perfect score: 126

Sequence: 1 TDLDKRGDNDISPFSGDGPFPKD 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /Genesec_101002:*

1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*

2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*

3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*

4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*

5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*

6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*

7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*

8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*

9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*

10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*

11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*

12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*

13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*

14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*

15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*

16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*

17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*

18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*

19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*

20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*

21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*

22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	126	100.0	23	AAU93726	Dental product bon
2	126	100.0	23	AAO20379	C-terminal amide
3	126	100.0	23	AAU93681	Dental product bon
4	126	100.0	97	AAO20331	Protein of matrix
5	126	100.0	430	AA153812	Amino acid sequenc
6	126	100.0	22	AAAB2669	Truncated phosphat
7	126	100.0	509	AAAB2920	Human osteoregulin
8	126	100.0	22	AAAB2920	Human osteoregulin
9	126	100.0	525	AAAB2669	Human osteoregulin
10	126	100.0	525	AAAB2669	Human osteoregulin
			23	ABB08526	Human phosphatonin

11	126	100.0	525	23	AAE13227	Human oncogenic os
12	126	100.0	540	22	AAAB2923	Human osteoregulin
13	126	100.0	556	22	AAAB2921	Human osteoregulin
14	101	80.2	40	23	AAU93703	Dental product bon
15	101	80.2	40	23	AAU93703	Dental product bon
16	86	68.3	19	23	AAE13230	Protein of matrix
17	84	66.7	15	23	AAU93725	Human OOM1 calcium
18	84	66.7	15	23	AAO20378	Dental product bon
19	84	66.7	33	23	AAU93710	C-terminal amide
20	84	66.7	33	23	AAO20360	Dental product bon
21	82	65.1	15	23	AAU93724	Peptide of matrix
22	82	65.1	15	23	AAO20377	Dental product bon
23	81	64.3	435	22	AAAB2918	C-terminal amide
24	81	64.3	441	22	AAAB2919	Ret osteoregulin.
25	80	63.5	275	22	AAU93507	Mouse osteoregulin
26	75	59.5	15	23	AAU93723	Human polypeptide
27	75	59.5	15	23	AAU93726	Dental product bon
28	73.5	58.3	40	23	AAU93704	C-terminal amide
29	73.5	58.3	40	23	AAO20354	Dental product bon
30	72.5	57.5	30	23	AAU93706	Protein of matrix
31	72.5	57.5	30	23	AAO20356	Dental product bon
32	72.5	57.5	35	23	AAU93705	Peptide of matrix
33	72.5	57.5	35	23	AAO20355	Dental product bon
34	66	52.4	12	23	AAO20372	Peptide of a calci
35	57	45.2	40	23	AAU93712	Dental product bon
36	57	45.2	40	23	AAO20362	Protein of matrix
37	57	45.2	45	23	AAU93707	Dental product bon
38	57	45.2	45	23	AAO20357	Dental product bon
39	53	42.1	109	22	AAE10214	Protein of matrix
40	53	42.1	1216	23	AAE12860	Human bone marrow
41	52	41.3	19	23	AAE13229	Human porlin peptid
42	52	41.3	301	17	AAW92999	Rhodobacter capsul
43	52	41.3	301	22	AAE1873	Sequence of porin
44	49.5	38.3	793	22	ABB6973	Drosophila melanog
45	49	38.9	9	23	AAE13232	Human OOM1 calcium

ALIGNMENTS

RESULT 1

AAU93726 standard; peptide: 23 AA.

ID AAU93726;

AC AAU93726;

XX 02-JUL-2002 (first entry)

XX

XX

XX

DE Dental product bone growth enhancing peptide #46.

XX

XX

KW Dental product; toothpaste; mouthwash; dental floss; bone growth;

KW Integrin binding motif; RGD; skeletal disease; tooth;

KW alveolar bone growth; osteoblast; odontoblast; osteopathic.

XX

OS Synthetic.

XX

XX WO200213775-A1.

XX

XX 21-FEB-2002.

XX

XX

XX 09-AUG-2001; 2001WO-US25101.

XX

XX 16-AUG-2000; 2000US-225879P.

XX

XX (BIGB-) BIG BEAR BIO INC.

XX

XX Yoneda T, Nomizu M, Kumagai Y;

XX

XX WPI: 2002-329525/36.

XX

XX Dental product useful for treating skeletal diseases e.g. dental

XX diseases comprises a base material and a compound comprising specific

XX amino acid in a sequence containing the integrin binding motif

XX Claim 7, Page 21: 44pp; English.
 PS
 CC The present invention relates to dental products such as toothpastes,
 CC mouthwash and dental floss comprising a base material and a compound
 CC which promotes bone growth. Such compounds are peptide sequences
 CC comprising 10-50 amino acids and containing an integrin binding
 CC motif such as RGD in the D- or L- form, preferably the L-configuration.
 CC The peptides of the invention are useful for treating or preventing
 CC skeletal diseases such as dental disease. The peptides enhance tooth
 CC and/or alveolar bone growth on areas where deterioration has occurred,
 CC as well as the growth or recruiting of osteoblast or odontoblast cells
 CC on the surface of the new skeletal growth. AAU93681-AAU93726 represent
 CC bone growth enhancing peptides for use in dental products.
 CC
 SQ Sequence 23 AA:

Query Match 100.0%; Score 126; DB 23; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1.8e-12;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDQERGDNDISFSGDGPFKD 23
 ID 1 TDQERGDNDISFSGDGPFKD 23
 DB

RESULT 2
 AAO20379
 ID AAO20379 standard; peptide; 23 AA.

AC AAO20379;
 XX
 DT 31-MAY-2002 (first entry)

DE C-terminal amidated synthesised peptide D-00006.

XX Bone growth; RGD motif; integrin binding motif; calcium binding motif;
 KM glycosaminoglycan binding motif; bone loss; renal phosphate excretion;
 KM alveolar; teeth; odontoblast; osteoclast; dental tissue; skeletal loss;
 KM weakness; D00006.

XX Synthetic.

XX Key Location/Qualifiers
 FT Modified-site 23
 FT /note="C-terminal amide"

PN WO200214360-A1.

XX 21-FEB-2002.

PD 14-AUG-2001; 2001WO-US25542.

PF 16-AUG-2000; 2000US-0641034.

PR 19-MAR-2001; 2001US-0812485.

XX (BIGB-) BIG BEAR BIO INC.

XX Kumagai Y, Blacher RW, Yoneda T;

XX WPI: 2002-291971/33.

XX New peptide compound useful for reducing bone loss, is capable of
 PT enhancing bone growth, and comprises an integrin binding motif,
 PT glycosaminoglycan binding motif or a calcium binding motif
 XX
 XX Example 1; Page 15; 50pp; English.

XX The invention relates to a peptide compound capable of enhancing bone
 CC growth, and comprising 10-50 amino acids in a sequence, where the amino
 CC acids are in D- or L- conformation and the sequence comprises a motif
 CC selected from an integrin binding motif, a glycosaminoglycan binding
 CC motif and a calcium binding motif. The peptide of the invention is useful

CC for reducing bone loss and for reducing renal phosphate excretion in an
 CC individual. The peptide is useful for promoting regeneration of alveolar
 CC bone and/or teeth, and increases the number and activity of odontoblasts
 CC /osteoclasts that help form dental tissues. The peptide is also useful
 CC for treating or preventing a condition associated with skeletal loss or
 CC weakness. This sequence represents a C-terminal amidated synthesised
 CC peptide D-00006 of the invention.
 CC
 SQ Sequence 23 AA:

Query Match 100.0%; Score 126; DB 23; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1.8e-12;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDQERGDNDISFSGDGPFKD 23
 ID 1 TDQERGDNDISFSGDGPFKD 23
 DB

RESULT 3
 AAU93681
 ID AAU93681 standard; protein; 97 AA.

XX AAU93681;

DT 02-JUL-2002 (first entry)

DE Dental product bone growth enhancing peptide #1.

XX Dental product; toothpaste; mouthwash; dental floss; bone growth;
 KM integrin binding motif; RGD; skeletal disease; dental disease; tooth;
 KM alveolar bone growth; osteoblast; odontoblast; osteopathic.

XX Synthetic.

PN WO200213775-A1.

XX 21-FEB-2002.

PF 09-AUG-2001; 2001WO-US25101.

PR 16-AUG-2000; 2000US-225879P.

XX (BIGB-) BIG BEAR BIO INC.

XX Yoneda T, Nomizu M, Kumagai Y;

XX WPI: 2002-329525/36.

XX Dental product useful for treating skeletal diseases e.g. dental
 PT diseases comprises a base material and a compound comprising specific
 PT amino acid in a sequence containing the integrin binding motif
 XX
 XX Disclosure; Page 11; 44pp; English.

XX The present invention relates to dental products such as toothpastes,
 CC mouthwash and dental floss comprising a base material and a compound
 CC which promotes bone growth. Such compounds are peptide sequences
 CC comprising 10-50 amino acids and containing an integrin binding
 CC motif such as RGD in the D- or L- form, preferably the L-configuration.
 CC The peptides of the invention are useful for treating or preventing
 CC skeletal diseases such as dental disease. The peptides enhance tooth
 CC and/or alveolar bone growth on areas where deterioration has occurred,
 CC as well as the growth or recruiting of osteoblast or odontoblast cells
 CC on the surface of the new skeletal growth. AAU93681-AAU93726 represent
 CC bone growth enhancing peptides for use in dental products.
 CC
 SQ Sequence 97 AA:

Query Match 100.0%; Score 126; DB 23; Length 97;
 Best Local Similarity 100.0%; Pred. No. 9.1e-12;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDLOERGNDISPFGDGPFD 23
 XX |||||||||||||||||||
 DB 43 TDLOERGNDISPFGDGPFD 65

RESULT 4
 AAO20331
 ID AAO20331 standard; protein: 97 AA.

AC AAO20331;

DT 31-MAY-2002 (first entry)

DE Protein of matrix extracellular phosphoglycoprotein containing RGD #1.

KM Bone growth; RGD motif; integrin binding motif; calcium binding motif;
 KM glycosaminoglycan binding motif; bone loss; renal phosphate excretion;
 KM alveolar teeth; odontoblast; osteoclast; dental tissue; skeletal loss;
 KM weakness; matrix extracellular phosphoglycoprotein.

OS Unidentified.

PN MO20214360-A1.

PD 21-FEB-2002.

PE 14-AUG-2001; 2001WO-US25542.

PR 16-AUG-2000; 2000US-0641034.

PR 19-MAR-2001; 2001US-0812485.

PA (BIGB-) BIG BEAR BIO INC.

PI Kumagai Y, Blacher RW, Yoneda T;

DR WPI: 2002-291971/33.

PT New peptide compound useful for reducing bone loss, is capable of
 PT enhancing bone growth, and comprises an integrin binding motif,
 PT glycosaminoglycan binding motif or a calcium binding motif

PS Disclosure; Page 11; 50pp; English.

XX The invention relates to a peptide compound capable of enhancing bone
 CC growth, and comprising 10-50 amino acids in a sequence, where the amino
 CC acids are in D- or L- conformation and the sequence comprises a motif
 CC selected from an integrin binding motif, a glycosaminoglycan binding
 CC motif and a calcium binding motif. The peptide of the invention is useful
 CC for reducing bone loss and for reducing renal phosphate excretion in an
 CC individual. The peptide is useful for promoting regeneration of alveolar
 CC bone and/or teeth, and increases the number and activity of odontoblasts
 CC /osteoclasts that help form dental tissues. The peptide is also useful
 CC for treating or preventing a condition associated with skeletal loss or
 CC weakness. This sequence represents a protein of a matrix extracellular
 CC phosphoglycoprotein containing an RGD motif of the invention.

SO Sequence 97 AA;

Query Match 100.0%; Score 126; DB 23; Length 97;
 Best Local Similarity 100.0%; Pred. No. 9; 1e-12;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDLOERGNDISPFGDGPFD 23
 XX |||||||||||||||||||
 DB 43 TDLOERGNDISPFGDGPFD 65

RESULT 5

AAV53812

ID AAV53812 standard; Protein: 430 AA.

AC AAV53812;

XX

DT 22-FEB-2000 (first entry)

XX Amino acid sequence of a human phosphatonin polypeptide.

DE Human; phosphatonin; Metastatic-tumour Excreted Phosphaturic-Element;
 KM MEPE; Na+-dependent phosphate cotransport; Vitamin D metabolism;
 KM bone mineralisation; phosphate metabolism related disease;
 KM hyperphosphatemia; renal osteodystrophy; renal dialysis;
 KM secondary hyperparathyroidism; osteitis fibrosa cystica; hypercalcaemia;
 KM x-linked hypophosphatemic rickets; hereditary hypophosphatemic rickets;
 KM hypomineralised bone lesion; stunted growth; cystic fibrosis;
 KM oncogenic hypophosphatemics osteomalacia; renal phosphate leakage;
 KM renal osteodystrophy; osteoporosis; Vitamin D resistant rickets;
 KM end organ resistance; renal Fanconi syndrome; autosomal rickets;
 KM Paget's disease; kidney failure; renal tubular acidosis; spure.

OS Homo sapiens.

FN Key Location/Qualifiers

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

FT Modified-site

8..10 Location/Qualifiers
 /note- "protein kinase C phosphorylation site"
 8..11 /note- "Caseln kinase II phosphorylation site"
 16..21 /note- "myristoylation site"
 40..47 /note- "Tyrosine kinase phosphorylation site"
 77..79 /note- "protein kinase C phosphorylation site"
 118..120 /note- "protein kinase C phosphorylation site"
 119..224 /note- "protein kinase C phosphorylation site"
 139..142 /note- "myristoylation site"
 143..148 /note- "Caseln kinase II phosphorylation site"
 152..154 /note- "myristoylation site"
 161..165 /note- "cell attachment tripeptide"
 177..180 /note- "glycosaminoglycan attachment site"
 /note- "Caseln kinase II phosphorylation site"
 194..197 /note- "Caseln kinase II phosphorylation site"
 199..202 /note- "Caseln kinase II phosphorylation site"
 203..205 /note- "Caseln kinase II phosphorylation site"
 224..227 /note- "protein kinase C phosphorylation site"
 228..231 /note- "Caseln kinase II phosphorylation site"
 228..230 /note- "Caseln kinase II phosphorylation site"
 238..241 /note- "protein kinase C phosphorylation site"
 266..271 /note- "Caseln kinase II phosphorylation site"
 291..296 /note- "myristoylation site"
 311..313 /note- "myristoylation site"
 312..314 /note- "protein kinase C phosphorylation site"
 315..320 /note- "protein kinase C phosphorylation site"
 319..321 /note- "myristoylation site"
 325..328 /note- "protein kinase C phosphorylation site"
 370..373 /note- "Caseln kinase II phosphorylation site"
 /note- "amidation site"

FT	Modified-site	382..386	/note- "Asu-glycosylation site"
FT	Modified-site	383..387	/note- "Asu-glycosylation site"
FT	Modified-site	384..386	/note- "protein kinase C phosphorylation site"
FT	Modified-site	389..394	/note- "myristoylation site"
FT	Modified-site	403..405	/note- "protein kinase C phosphorylation site"
FT	Modified-site	405..408	/note- "CAMP and CGMP dependent protien kinase phosphorylation site"
FT	Modified-site	408..410	/note- "protein kinase C phosphorylation site"
FT	Modified-site	409..411	/note- "protein kinase C phosphorylation site"
FT	Modified-site	423..426	/note- "Caseln kinase II phosphorylation site"
FT	Modified-site	425..428	/note- "Caseln kinase II phosphorylation site"
FT	Modified-site	427..430	/note- "Caseln kinase II phosphorylation site"
FT	Modified-site		

PN	WO9960017-A2.
XX	
XX	
PD	25-NOV-1999.
XX	
XX	
PF	18-MAY-1999; 99WO-EP03403.
XX	
XX	
PR	18-MAY-1998; 98GB-0010681.
PR	04-SEP-1998; 98GB-0019387.

PA (UNLO) UNIV COLLEGE LONDON.

DR - WPI; 2000-053262/04.
DR N-PSDB; AAZ36447.

PT New polypeptides involved in the regulation of phosphate metabolism
PT useful for diagnosing and treating disorders related to phosphate metabolism -

PS Claim 6; Fig 8; 136pp; English.

The present sequence represents a phosphatonin/polypeptide (also called Metastatic-tumour Excreted Phosphaturic-Element (MEPE)). The level of phosphatonin in a subject modulates Na⁺-dependent phosphate cotransport vitamin D metabolism and/or bone mineralisation. The phosphatonin polypeptides, polynucleotides, vectors and antibodies are used to treat phosphate metabolism related disease. They are used for treatment of hyperphosphatemia, or renal osteodystrophy, hyperphosphatemia in renal dialysis/pre-dialysis, secondary hyperparathyroidism or osteitis fibrosa cystica, or x-linked hypophosphatemic rickets, hereditary hypophosphatemic rickets with hypocalcaemia (HNHR), hypomieralised bone lesions, stunted growth in juveniles, oncogenic hypophosphatemic osteomalacia, renal phosphate leakage, renal osteodystrophy, osteoporosis, vitamin D resistant rickets, end organ resistance, renal Fanconi syndrome, autosomal rickets, Paget's disease, kidney failure, renal tubular acidosis, cystic fibrosis or sprue. The polypeptide may also be used to manufacture combined preparations for simultaneous, separate or sequential use for the treatment of phosphate metabolism disorders. A transformed osteoblast or bone cell line capable of phosphatonin overexpression is useful for the production of phosphatonin.

SQ Sequence 430 AA;

Query Match	100.0%;	Score 126;	DB 21;	Length 430;
Best Local Similarity	100.0%;	Pred. No. 5e-11;		
Matches	23;	Conservative	0;	Mismatches 0; Indels 0

```
QY      1  TDLQERGNDISPFGDGQPKD  23
          |||||
Db      147 TDLQERGNDISPFGDGQPKD  169
```

RESULT 6
AAB62669
ID AAB62669 standard; Protein; 430 AA

AC AAB62669

DT 23-JUL-2001 (first entry)

Truncated phosphatoin polypeptide (truncated MEPE).

KM Metastatic-tumour excreted phosphatonic element; MEPE, phosphatonin;;
 KM phosphate; vitamin-D; skeletal formation; mineralization; truncated;
 KM osteopathic; antigen; cytostatic; human.

OS Homo sapiens

FH	Key	Location/Qualifiers
FM	W43-44660000	151 103

```

FT      /note="the amino acid residues in this region are
FT      indicated incorrectly in the sequence provided
FT      in the sequence listing"

```

PN W0200132878-A2.

PD 10-MAY-2001

31-OCT-2000; 2000WO-EPI0747.

PR	04-NOV-1999;	9905-0434185
PR	08-NOV-1999;	00CE-0035434

PA (UNLO) UNIV COLLEGE LONDON.

PI ROWE P;

DR WPI; 2001-343487/36.

DR N-PSDB; AAF83759.

PT New phosphatonin polypeptide a regulator of phosphate metabolism, for
PT diagnosing and treating disorders of phosphate, vitamin-D metabolism
PT skeletal formation e.g. osteoporosis, Paget's disease, gout -

PS Claim 4; Fig 8; 135pp; English.

The invention relates to a novel human protein, metastatic-tumour excreted phosphatonic element (MPE) or phosphatonin (modulator of phosphate and vitamin-D metabolism). The phosphatonin (modulator of polynucleotides and specific antihodors are useful for treating a disorder of phosphate or vitamin D metabolism, skeletal formation and mineralization. Phosphatonins are used to treat hyperphosphatemia, renal osteodystrophy, secondary hyperparathyroidism, osteitis fibrosa cystica or gout. It is used to prepare a medicament for treating x-linked hypophosphatemic rickets, hereditary hypophosphatemic rickets with hypocalcemia (HHRH), hypomineralized bone lesions, stunted growth in juveniles, oncogenic hypophosphatemic osteomalacia, renal phosphate leakage, renal osteodystrophy, osteoporosis, vitamin D-resistant rickets, end organ resistance, renal Fanconi syndrome, autosomal rickets, Paget's disease, kidney failure, renal tubular acidosis, cystic fibrosis or sprue. Phosphatonin polynucleotides are useful as molecular weight markers on Southern gels, as diagnostic probes for detecting the presence of a specific mRNA. Phosphatonin polypeptides are also useful for identifying agonists and antagonists, compounds which bind to phosphatonin and drug candidates for therapy of phosphate metabolism disorders. The present sequence represents a truncated form of phosphatonin (MPE).

SQ Sequence 430 AA;

Query Match 100.0%; Score 126; DB 22; Length 430;
 Best Local Similarity 100.0%; Pred. No. 5e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDLOERGDNDISPFSGDQPFKD 23
 147 TDLOERGDNDISPFSGDQPFKD 169

RESULT 7
 AAB82922

ID AAB82922 standard; Protein: 509 AA.

AC AAB82922;

DT 21-DEC-2001 (first entry)

DE Human osteoregulin (mature polypeptide).

KW Osteoregulin; human; bone; homeostasis; adipose; calcification;
 KW atherosclerosis; osteoporosis; osteopathic; antiarteriosclerotic;
 KW therapy.

OS Homo sapiens.

PN EP1130098-A2.

PD 05-SEP-2001.

PF 27-FEB-2001; 2001EP-0301768.

PR 29-FEB-2000; 2000US-185617P.

PR 22-SEP-2000; 2000US-234500P.

PA (PF12) PFIZER PROD INC.

PI Brown TA, De Wet JR, Gowen LC, Hames LM;

DR WPI; 2001-604111/69.

DR N-PSDB; AAH26810.

PT Novel osteoregulin polypeptide useful for regulating bone homeostasis,
 PT adiposity and calcification of atherosclerotic plaques comprises
 PT measuring the activity of osteoregulin -
 PS Claim 1; Page 54-55; 90pp; English.

CC The present sequence is that of human osteoregulin mature
 CC polypeptide, i.e. lacking an N-terminal signal sequence.
 CC Osteoregulin is a novel protein which plays a role in regulating
 CC bone homeostasis, adiposity, and the calcification of
 CC atherosclerotic plaques. 2 splice variants of human osteoregulin
 CC were identified (see also AAB82922). The invention provides novel
 CC osteoregulin proteins, nucleic acids which encode them, vectors,
 CC antibodies, host cells which express heterologous osteoregulins, and
 CC animal cells and mammals with a targeted disruption of an
 CC osteoregulin gene. The invention also provides screening assays
 CC to identify modulators of osteoregulin activity as well as methods
 CC of treating mammals for diseases or disorders associated with
 CC osteoregulin activity. The modulators of activity may be useful
 CC in the manufacture of a medicament for, as well as for treating, a
 CC mammal in need of regulation of bone mass and/or density, adiposity,
 CC vascular flexibility, and/or atherosclerotic plaque calcification
 CC (claimed), for treating and preventing osteoporosis, and for
 CC stimulating bone repair and regeneration.

SQ Sequence 509 AA;

Query Match 100.0%; Score 126; DB 22; Length 509;
 Best Local Similarity 100.0%; Pred. No. 6.1e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDLOERGDNDISPFSGDQPFKD 23

Db 226 TDLOERGDNDISPFSGDQPFKD 248

RESULT 8
 AAB82920
 ID AAB82920 standard; Protein: 525 AA.

AC AAB82920;

DT 21-DEC-2001 (first entry)

DE Human osteoregulin.

KW Osteoregulin; human; bone; homeostasis; adipose; calcification;
 KW atherosclerosis; osteoporosis; osteopathic; antiarteriosclerotic;
 KW therapy.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Peptide 1..37 /label= signal_peptide

FT Protein 38..525 /label= Mature_protein

PN EP1130098-A2.

PD 05-SEP-2001.

PF 27-FEB-2001; 2001EP-0301768.

PR 29-FEB-2000; 2000US-185617P.

PR 22-SEP-2000; 2000US-234500P.

PA (PF12) PFIZER PROD INC.

PI Brown TA, De Wet JR, Gowen LC, Hames LM;

DR WPI; 2001-604111/69.

DR N-PSDB; AAH26808.

PT Novel osteoregulin polypeptide useful for regulating bone homeostasis,
 PT adiposity and calcification of atherosclerotic plaques comprises
 PT measuring the activity of osteoregulin -
 PS Claim 1; Page 45-47; 90pp; English.

CC The present sequence is that of human osteoregulin, a novel protein
 CC which plays a role in regulating bone homeostasis, adiposity, and
 CC the calcification of atherosclerotic plaques. The sequence is
 CC predicted from the nucleotide sequence (see AAH26808) of isolated
 CC osteoblast cDNA. A splice variant of human osteoregulin was also
 CC identified (see AAB82921). The invention provides novel
 CC osteoregulin proteins, nucleic acids which encode them, vectors,
 CC antibodies, host cells which express heterologous osteoregulins, and
 CC animal cells and mammals with a targeted disruption of an
 CC osteoregulin gene. The invention also provides screening assays
 CC to identify modulators of osteoregulin activity as well as methods
 CC of treating mammals for diseases or disorders associated with
 CC osteoregulin activity. The modulators of activity may be useful
 CC in the manufacture of a medicament for, as well as for treating, a
 CC mammal in need of regulation of bone mass and/or density, adiposity,
 CC vascular flexibility, and/or atherosclerotic plaque calcification
 CC (claimed), for treating and preventing osteoporosis, and for
 CC stimulating bone repair and regeneration.

SQ Sequence 525 AA;

Query Match 100.0%; Score 126; DB 22; Length 525;
 Best Local Similarity 100.0%; Pred. No. 6.3e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDLQERGDNDISFSGDGPFD 23
 DB 242 TDLQERGDNDISFSGDGPFD 264

RESULT 9

ID AAB62689 standard; Protein: 525 AA.

AC AAB62689;

DT 23-JUL-2001 (first entry)

DE Phosphatonin polypeptide (MEPE).

KM Metastatic tumour excreted phosphaturic element; MEPE; phosphatonin;
 KM phosphate; vitamin-D; skeletal formation; mineralization; truncated;
 KM osteopachic; antlygout; cytostatic; human.

OS Homo sapiens.

PN MO200132878-A2.

PD 10-MAY-2001.

PF 31-OCT-2000; 2000WO-EP10747.

PR 04-NOV-1999; 99US-0434185.

PR 08-NOV-1999; 99GB-0026424.

PA (UNLO) UNIV COLLEGE LONDON.

PI Rowe P;

DR WPI; 2001-343487/36.

DR N-PSDB; AAF83764.

PT New phosphatonin polypeptide a regulator of phosphate metabolism, for
 PT diagnosing and treating disorders of phosphate, vitamin-D metabolism,
 PT skeletal formation e.g. osteoporosis, Paget's disease, gout

PS Claim 4: Page 133-134; 135pp; English.

XX The invention relates to a novel human protein, metastatic-tumour
 CC excreted phosphaturic element (MEPE) or phosphatonin (modulator of
 CC phosphate and vitamin-D metabolism). The phosphatonin polypeptides,
 CC polynucleotides and specific antibodies are useful for treating a
 CC disorder of phosphate or vitamin D metabolism, skeletal formation and
 CC mineralization. Phosphatonins are used to treat hyperphosphatemia, renal
 CC osteodystrophy, secondary hyperparathyroidism, osteitis fibrosa cystica
 CC or gout. It is used to prepare a medicament for treating x-linked
 CC hypophosphatemic rickets, hereditary hypophosphatemic rickets with
 CC hypercalcaemia (HHRH), hypomineralized bone lesions, stunted growth in
 CC juveniles, oncogenic hypophosphatemic osteomalacia, renal phosphate
 CC leakage, renal osteodystrophy, osteoporosis, vitamin-D resistant rickets,
 CC end organ resistance, renal Fanconi syndrome, autosomal rickets, Paget's
 CC disease, kidney failure, renal tubular acidosis, cystic fibrosis or
 CC sprue. Phosphatonin polynucleotides are useful as molecular weight
 CC markers on Southern gels, as diagnostic probes for detecting the presence
 CC of a specific mRNA. Phosphatonin polypeptides are also useful for
 CC identifying agonists and antagonists, compounds which bind to
 CC phosphatonin and drug candidates for therapy of phosphate metabolism
 CC disorders. The present sequence represents the amino acid sequence of
 CC the entire phosphatonin (MEPE).

SQ Sequence 525 AA;

Query Match 100.0%; Score 126; DB 22; Length 525;

Best Local Similarity 100.0%; Pred. No. 6.3e-11;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDLQERGDNDISFSGDGPFD 23
 DB 242 TDLQERGDNDISFSGDGPFD 264

DB 242 TDLQERGDNDISFSGDGPFD 264

RESULT 10

ID ABB08526 standard; Protein: 525 AA.

AC ABB08526;

DT 23-MAY-2002 (first entry)

DE Human phosphatonin.

KM Human phosphatonin; cytostatic; antidiabetic; antiinflammatory;
 KM hyperphosphemia; arteriosclerosis; heart failure;
 KM diabetic renal disease; kidney failure; cystic fibrosis.

OS Homo sapiens.

PN MO200198495-A1.

PD 27-DEC-2001.

PF 20-JUN-2001; 2001WO-JP05263.

PR 21-JUN-2000; 2000JP-0191088.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Kurokawa T, Yamada T, Morimoto S;

DR WPI; 2002-139791/18.

DR N-PSDB; ABA99159, ABA99160.

PT Phosphatonin of human origin and DNA encoding it for diagnosis and
 PT treatment of diseases associated with disorders of phosphate
 PT metabolism, e.g., hyperphosphemia, arteriosclerosis, heart failure,
 PT diabetic renal disease and kidney failure

PS Claim 1: Fig 1-4; 130pp; Japanese.

XX This invention relates to human phosphatonin which has
 CC the activity of lowering blood phosphate and increasing urinary
 CC phosphate. The proteins are cytostatic, antidiabetic and
 CC antiinflammatory in their action. Phosphatonin down-regulates
 CC sodium-dependent phosphate transport in the kidney, it down-regulates
 CC 25-hydroxy-vitamin D3-1alpha-hydroxylase in the kidney and up-regulates
 CC 25-hydroxy-vitamin D3-24-hydroxylase in the kidney. Phosphatonin is
 CC useful in the diagnosis, treatment and prevention of phosphate
 CC metabolism related diseases such as hyperphosphemia, arteriosclerosis,
 CC heart failure, diabetic renal disease, kidney failure, acute coronary
 CC disease and cystic fibrosis. This sequence represents human
 CC phosphatonin.

SQ Sequence 525 AA;

Query Match 100.0%; Score 126; DB 23; Length 525;

Best Local Similarity 100.0%; Pred. No. 6.3e-11;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDLQERGDNDISFSGDGPFD 23
 DB 242 TDLQERGDNDISFSGDGPFD 264

RESULT 11

ID AAE13227 standard; Protein: 525 AA.

AC AAE13227;

DT 12-FEB-2002 (first entry)

DE	Human oncogenic osteomalacia-related protein-1 (OOM-1).
KW	Human; oncogenic osteomalacia-related protein 1; OOM1; gene therapy;
KM	phosphate homeostasis-related disease; rickets; osteomalacia; cadant;
KW	rhabdomyolysis; cardiomyopathy; tumoural calcinosis; renal failure;
KM	bone mineralisation; phosphaturia; cellular process.
XX	
OS	Homo sapiens.
FH	
FH	Key
FT	Peptide
FT	/label= Signal_peptide
FT	Protein
FT	17..525
FT	(note= "Mature oncogenic osteomalacia-related protein-1
FT	"OOM-1")
FT	130..142
FT	/note= "Calcium binding motif"
FT	235..258
FT	/note= "Calcium binding motif"
FT	Domain
FT	Region
FT	247..249
FT	/note= "Calcium binding motif"
FT	264..275
FT	/note= "Cell attachment site"
FT	Domain
FT	412..424
FT	/note= "Calcium binding motif"
FT	Modified-site
FT	442..455
FT	/note= "Potential PKA phosphorylation site"
FT	Modified-site
FT	477..481
FT	/note= "Glycosylation site"
FT	Domain
FT	500..503
FT	/note= "Bacterial extracellular solute-binding
FT	protein motif"
PN	
PD	WO200172826-A2.
PD	04-OCT-2001.
PE	
PR	22-MAR-2001; 2001WO-US09289.
PR	24-MAR-2000; 2000US-191786P.
PR	19-OCT-2000; 2000US-241598P.
PA	(GENZ) GENZYME CORP.
PA	(UTJO) UNIV JOHN HOPKINS.
PI	Schlavi S, Madden S, Manavalan P, Levine MD, Jan De Beur S;
PT	WT: 2002-010720/O1.
PT	New polynucleotide encoding a polypeptide useful for identifying
PT	polynucleotide expression or polypeptide activity modulators used for
PT	treating oncogenic osteomalacia, comprises the oncogenic
PT	osteomalacia-related gene -
PS	Claim 2; Fig 2; 65pp; English.
XX	
XX	The invention relates to oncogenic osteomalacia-related protein 1
CC	designated OOM1 (oncogenic osteomalacia-related factor) and its DNA
CC	molecule. OOM1 protein is useful for treating a phosphate homeostasis-
CC	related disease such as X-linked hypophosphataemia rickets, oncogenic
CC	osteomalacia, rhabdomyolysis, cardiomyopathy, tumoural calcinosis and
CC	renal failure. OOM1 proteins and DNA's are useful for modulating the
CC	phenotype of a neoplastic cell associated with oncogenic osteomalacia;
CC	modulating bone mineralisation; treating a disease characterised by
CC	abnormal bone mineralisation; and treating an oncogenic osteomala-
CC	associated symptom such as hypophosphataemia, phosphaturia, low serum
CC	concentrations of 1,25-dihydroxy vitamin D or osteomalacia. OOM1 is
CC	useful for modulating renal phosphate transport, which involves altering
CC	the activity of OOM1 within a cell and then administering serine protease
CC	11 to the cell. OOM1 DNA is useful for generating non-human transgenic
CC	animal models, and for searching and identifying single nucleotide
CC	polymorphisms which are mutants, variants of the gene in human
CC	population. OOM1 is useful as an immunogen to produce antibodies against

CC OOM1 and in vitro assays to screen for agents or drugs, which modulate
CC cellular processes. The present sequence is human oncogenic
CC osteomalacia-related protein-1 (OOM-1).
CC Note: The present sequence states that this sequence is encoded by the
CC DNA sequence shown in SEQ ID NO: 1 (AAD21889) of the specification.
CC However this does not appear to be the case.
XX

XX Sequence 525 AA:

XX

XX Query Match 100.0%; Score 126; DB 23; Length 525;
XX Best Local Similarity 100.0%; Pred. No. 6,3e-11;
XX Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

XX

XX 1 TDQERGDNDISPFSGDGPFRD 23
XX |||||||||||||||||||
XX TDQERGDNDISPFSGDGPFRD 264

XX

XX DB 242 TDQERGDNDISPFSGDGPFRD 264

XX

XX RESULT 12

XX AAB82923

XX ID AAB82923 standard; Protein: 540 AA.

XX AC AAB82923;

XX XX

XX DT 21-DEC-2001 (first entry)

XX XX

XX DE Human osteoregulin (mature polypeptide).

XX XX

XX KW Osteoregulin; human; bone; homeostasis; adipose; calcification;
XX KW atherosclerosis; osteoporosis; osteopathic; antiarteriosclerotic;
XX KW therapy.

XX XX

XX OS Homo sapiens.

XX PN EPI130098-A2.

XX PD 05-SEP-2001.

XX XX

XX PF 27-FEB-2001; 2001EP-0301768.

XX PR 29-FEB-2000; 2000US-185617P.

XX PR 22-SEP-2000; 2000US-234500P.

XX XX

XX PA (PE12) PFIZER PROD INC.

XX XX

XX PI Brown TA, De Wet JR, Gowen LC, Hames LM;

XX DR WPI; 2001-604111/69.

XX DR N-PSDB; AAB26811.

XX XX

XX PT Novel osteoregulin polypeptide useful for regulating bone homeostasis,
XX PT adiposity and calcification of atherosclerotic plaques comprises
XX PT measuring the activity of osteoregulin .

XX

XX Claim 1; Page 59-61; 90pp; English.

XX

XX The present sequence is that of human osteoregulin mature
XX polypeptide, i.e. lacking an N-terminal signal sequence.
XX Osteoregulin is a novel protein in which plays a role in regulating
XX bone homeostasis, adiposity, and the calcification of
XX atherosclerotic plaques. 2 Splice variants of human osteoregulin
XX were identified (see also AAB82922). The invention provides novel
XX osteoregulin proteins, nucleic acids which encode them, vectors,
XX antibodies, host cells which express heterologous osteoregulins, and
XX animal cells and mammals with a targeted disruption of an
XX osteoregulin gene. The invention also provides screening assays
XX to identify modulators of osteoregulin activity as well as methods
XX of treating mammals for diseases or disorders associated with
XX osteoregulin activity. The modulators of activity may be useful
XX in the manufacture of a medicament for, as well as for treating, a
XX mammal in need of regulation of bone mass and/or density, adiposity,
XX vascular flexibility, and/or atherosclerotic plaque calcification
XX (claimed), for treating and preventing osteoporosis, and for

CC stimulating bone repair and regeneration.
 XX Sequence 540 AA;
 Query Match 100.0%; Score 126; DB 22; Length 540;
 Best Local Similarity 100.0%; Pred. No. 6.5e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 TDLOERGDNDISPFSGDQPFKD 23
 Db 257 TDLOERGDNDISPFSGDQPFKD 279

RESULT 13
 AAB82921
 ID AAB82921 standard; Protein; 556 AA.
 XX AAB82921;
 XX 21-DEC-2001 (first entry)
 XX Human osteoregulin.
 XX
 XX Osteoregulin; human; bone; homeostasis; adipose; calcification;
 KM atherosclerosis; osteoporosis; osteopathic; antiarteriosclerotic;
 KM therapy.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH Peptide 1..37
 FT /label= Signal_peptide
 FT Protein 38..556
 FT /label= Mature_protein
 FT Misc-difference 57
 FT /note= "encoded by TCA"
 FT
 XX EPI130098-A2.
 PN
 PD 05-SEP-2001.
 PD
 PF 27-FEB-2001; 2001EP-0301768.
 PF
 PR 29-FEB-2000; 2000US-185617P.
 PR 22-SEP-2000; 2000US-234500P.
 XX
 PA (PEIZ) PFIZER PROD INC.
 XX
 PI Brown TA, De Wet JR, Gowen LC, Hames LM;
 XX
 DR MPI: 2001-604111/59.
 DR N-PSDB; AAH26809.
 XX
 PT Novel osteoregulin polypeptide useful for regulating bone homeostasis,
 PT adiposity and calcification of atherosclerotic plaques comprises
 PT measuring the activity of osteoregulin .
 XX
 PS Claim 1; Page 48-49; 90pp; English.
 XX
 XX The present sequence is that of human osteoregulin, a novel protein
 CC which plays a role in regulating bone homeostasis, adiposity, and
 CC the calcification of atherosclerotic plaques. The sequence is
 CC predicted from the nucleotide sequence (see AAH26809) of isolated
 CC osteoblast cDNA. A splice variant of human osteoregulin was also
 CC identified (see AAB82920). The invention provides novel
 CC osteoregulin proteins, nucleic acids which encode them, vectors,
 CC antibodies, host cells which express heterologous osteoregulins, and
 CC animal cells and mammals with a targeted disruption of an
 CC osteoregulin gene. The invention also provides screening assays
 CC to identify modulators of osteoregulin activity as well as methods
 CC of treating mammals for diseases or disorders associated with
 CC osteoregulin activity. The modulators of activity may be useful
 CC in the manufacture of a medicament for, as well as for treating, a

CC mammal in need of regulation of bone mass and/or density, adiposity,
 CC vascular flexibility, and/or atherosclerotic plaque calcification
 CC (claimed), for treating and preventing osteoporosis, and for
 CC stimulating bone repair and regeneration.
 XX Sequence 556 AA;
 Query Match 100.0%; Score 126; DB 22; Length 556;
 Best Local Similarity 100.0%; Pred. No. 6.7e-11;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 TDLOERGDNDISPFSGDQPFKD 23
 Db 273 TDLOERGDNDISPFSGDQPFKD 295

RESULT 14
 AAU93703
 ID AAU93703 standard; peptide; 40 AA.
 XX AAU93703;
 AC
 XX 02-JUL-2002 (first entry)
 DT
 XX Dental product bone growth enhancing peptide #23.
 DE
 XX Dental product: toothpaste; mouthwash; dental floss; bone growth;
 KM integrin binding motif; RGD; skeletal disease; dental disease; tooth;
 KM alveolar bone growth; osteoblast; odontoblast; osteopathic.
 XX
 XX Synthetic.
 OS
 XX PN NC0200213775-A1.
 PN
 PD 21-FEB-2002.
 PD
 PF 09-AUG-2001; 2001MO-US25101.
 PF
 PR 16-AUG-2000; 2000US-225879P.
 PR
 XX (BIGB-) BIG BEAR BIO INC.
 PA
 PI Yoneda T, Nomizu M, Kumagai Y;
 XX
 DR WPI; 2002-329525/36.
 DR
 PT Dental product useful for treating skeletal diseases e.g. dental
 PT diseases comprises a base material and a compound comprising specific
 PT amino acid in a sequence containing the integrin binding motif .
 XX
 PS Claim 6; Page 20; 44pp; English.
 XX
 XX The present invention relates to dental products such as toothpastes,
 CC mouthwash and dental floss comprising a base material and a compound
 CC which promotes bone growth. Such compounds are peptide sequences
 CC comprising 10-50 amino acids and containing an integrin binding
 CC motif such as RGD in the D- or L- form, preferably the L-configuration.
 CC The peptides of the invention are useful for treating or preventing
 CC skeletal diseases such as dental disease. The peptides enhance tooth
 CC and/or alveolar bone growth on areas where deterioration has occurred,
 CC as well as the growth or recruiting of osteoblast or odontoblast cells
 CC on the surface of the new skeletal growth. AAU93681-AAU93726 represent
 CC bone growth enhancing peptides for use in dental products.
 XX
 XX Sequence 40 AA;
 Query Match 80.2%; Score 101; DB 23; Length 40;
 Best Local Similarity 100.0%; Pred. No. 2.4e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 6 RGDNDISPFSGDQPFKD 23
 Db 1 RGDNDISPFSGDQPFKD 18

RESULT 15

AAO20353 standard: protein; 40 AA.

AAO20353:

31-MAY-2002 (first entry)

Protein of matrix extracellular phosphoglycoprotein containing RGD #23.

Bone growth; RGD motif; integrin binding motif; calcium binding motif;
 glycosaminoglycan binding motif; bone loss; renal phosphate excretion;
 alveolar; teeth; odontoblast; osteoclast; dental tissue; skeletal loss;
 weakness; matrix extracellular phosphoglycoprotein.

Unidentified.

MO200214360-A1.

21-FEB-2002.

14-AUG-2001: 2001WO-US25542.

16-AUG-2000: 2000US-0641034.

19-MAR-2001: 2001US-0812485.

(BIGB-) BIG BEAR BIO INC.

Kumagai Y, Blacher RW, Yoneda T;

WPI: 2002-291971/33.

New peptide compound useful for reducing bone loss, is capable of
 enhancing bone growth, and comprises an integrin binding motif,
 glycosaminoglycan binding motif or a calcium binding motif.

Disclosure: Page 12: 50pp; English.

The invention relates to a peptide compound capable of enhancing bone
 growth, and comprising 10-50 amino acids in a sequence, where the amino
 acids are in D- or L- conformation and the sequence comprises a motif
 selected from an integrin binding motif, a glycosaminoglycan binding
 motif and a calcium binding motif. The peptide of the invention is useful
 for reducing bone loss and for reducing renal phosphate excretion in an
 individual. The peptide is useful for promoting regeneration of alveolar
 bone and/or teeth, and increases the number and activity of alveolar
 osteoclasts that help form dental tissues. The peptide is also useful
 for treating or preventing a condition associated with skeletal loss or
 weakness. This sequence represents a protein of a matrix extracellular
 phosphoglycoprotein containing an RGD motif of the invention.

Sequence 40 AA:

Query Match 80.2%; Score 101; DB 23; Length 40;

Best Local Similarity 100.0%; Pred. No. 2.4e-08;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 RGDNDISPFSGDGPFPKD 23

DB 1 RGDNDISPFSGDGPFPKD 18

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 13:16:42 : Search time 14 Seconds
(without alignments)
48.338 Million cell updates/sec

Title: US-09-641-034-47

Perfect score: 126

Sequence: 1 IDLQERGDNDISFGDQPFKD 23

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 262574 segs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA: *
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/Backfilest1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	41.3	301	2	US-08-355-844-2
2	52	41.3	301	5	PCT-US95-16126-2
3	47	37.3	815	4	US-09-134-001C-3614
4	47	37.3	1198	4	US-09-199-637A-405
5	45.5	36.1	96	2	US-08-341-843B-9
6	45.5	36.1	96	2	US-08-427-497E-14
7	45.5	36.1	1260	4	US-08-506-296B-21
8	44	34.9	348	1	US-08-176-126B-2
9	44	34.9	348	1	US-08-669-435-2
10	44	34.9	348	5	PCT-US94-14431A-2
11	44	34.9	475	1	US-08-278-091-6
12	44	34.9	475	1	US-08-483-859-6
13	44	34.9	475	1	US-08-472-173-6
14	44	34.9	475	1	US-08-350-741-2
15	44	34.9	475	1	US-08-487-167-6
16	44	34.9	475	2	US-08-482-816-6
17	44	34.9	475	2	US-08-296-149-6
18	44	34.9	475	2	US-08-801-499-6
19	44	34.9	475	2	US-08-463-875A-2
20	44	34.9	475	2	US-08-615-271-6
21	44	34.9	475	3	US-09-074-660-6
22	44	34.9	475	3	US-09-074-659-6
23	44	34.9	475	4	US-09-106-468-6
24	44	34.9	475	4	US-09-106-466A-6
25	44	34.9	475	4	US-09-106-467-6
26	43	34.1	169	3	US-08-476-509B-28
27	42.5	33.7	306	1	US-08-330-978-1

28	42.5	33.7	306	1	US-08-474-042-1	Sequence 1, Appl 1
29	42.5	33.7	306	1	US-08-484-558-1	Sequence 1, Appl 1
30	42.5	33.7	306	1	US-08-774-592-1	Sequence 1, Appl 1
31	42.5	33.7	448	1	US-08-295-411-3	Sequence 3, Appl 1
32	42.5	33.7	448	2	US-08-955-471-3	Sequence 3, Appl 1
33	42.5	33.7	448	5	PCT-US92-10068-1	Sequence 1, Appl 1
34	42.5	33.7	448	5	PCT-US92-10242-3	Sequence 1, Appl 1
35	42.5	33.7	488	1	US-08-487-037-1	Sequence 1, Appl 1
36	42.5	33.7	545	2	US-08-990-114-1	Sequence 1, Appl 1
37	42.5	33.7	545	4	US-09-241-333-1	Sequence 1, Appl 1
38	41.5	32.9	364	4	US-08-858-207A-313	Sequence 13, Appl 1
39	41.5	32.9	811	4	US-08-983-275-2	Sequence 2, Appl 1
40	40.5	32.1	397	4	US-09-087-134-17	Sequence 17, Appl 1
41	40	31.7	469	2	US-08-416-870C-6	Sequence 6, Appl 1
42	40	31.7	795	4	US-09-370-807-6	Sequence 6, Appl 1
43	40	31.7	795	4	US-09-921-259-6	Sequence 6, Appl 1
44	40	31.7	831	2	US-09-047-026A-4	Sequence 4, Appl 1
45	40	31.7	912	5	PCT-US95-03747-2	Sequence 2, Appl 1

ALIGNMENTS

RESULT 1
US-08-355-844-2
Sequence 2, Application US/08355844
Patent No. 5940307
GENERAL INFORMATION:
APPLICANT: Fischbarg, Jorge
APPLICANT: Czelegiedy, Ferenc
APPLICANT: Isevorich, Pavel
APPLICANT: Li, Jun
TITLE OF INVENTION: A METHOD FOR PREDICTING PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/355,844
FILING DATE: 14-DEC-1994
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Tang, Henry Y.S.
REGISTRATION NUMBER: 29,705
REFERENCE/DOCKET NUMBER: A29927-50/29910
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-765-2519
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 301 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORGANISM: Rhodobacter capsulatus
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..301
OTHER INFORMATION: Porin protein
US-08-355-844-2

Query Match 41.3%; Score 52; DB 2; Length 301;
Best Local Similarity 58.8%; Pred. No. 2.8;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 TDLQERGDNDISPSGD 17
|||:|||||:|||
Db 92 TDLDRGNDIPYLTGD 108

RESULT 2

PCT-US95-16126-2
Sequence 2, Application PC/TUS9516126
GENERAL INFORMATION:
APPLICANT: Fischbarg, Jorge
APPLICANT: Czelegedy, Ferenc
APPLICANT: Iserovich, Pavel
APPLICANT: Li, Jun
APPLICANT: Cheung, Min
TITLE OF INVENTION: A METHOD FOR PREDICTING PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSER: Brumbaugh, Graves, Donohue & Raymond
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10112-0228
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/16126
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/355,844
FILING DATE: 14-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Tang, Henry T.S.
REGISTRATION NUMBER: 29,705
REFERENCE/DOCKET NUMBER: A29927-50/29910
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-408-2586
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 301 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Rhodobacter capsulatus
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..301
OTHER INFORMATION: Porin protein
PCT-US95-16126-2

Query Match 41.3%; Score 52; DB 5; Length 301;
Best Local Similarity 58.8%; Pred. No. 2.8;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

2Y 1 TDLQERGDNDISPSGD 17
|||:|||||:|||
Db 92 TDLDRGNDIPYLTGD 108

RESULT 3

US-09-134-001C-3614
Sequence 3614, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3614
LENGTH: 815
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3614

Query Match 37.3%; Score 47; DB 4; Length 815;
Best Local Similarity 58.8%; Pred. No. 50;
Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 7 GDNDISPSGDCQPKD 23
|||:|||||:|||
Db 481 GDEDLPSVGPQGVFKD 497

RESULT 4

US-09-199-637A-405
Sequence 405, Application US/09199637A
Patent No. 6355411
GENERAL INFORMATION:
APPLICANT: Ausubel, Frederick
APPLICANT: Goodman, Howard M.
APPLICANT: Rahne, Laurence G.
APPLICANT: Mahajan-Miklos, Shailina
APPLICANT: Tan, Man-Wah
APPLICANT: Drenkard, Eliana
APPLICANT: Tsongalis, John
TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
FILE REFERENCE: 00786/361002
CURRENT APPLICATION NUMBER: US/09/199,637A
PRIOR FILING DATE: 1998-11-25
PRIOR APPLICATION NUMBER: 60/066,517
PRIOR FILING DATE: 1997-11-25
NUMBER OF SEQ ID NOS: 437
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 405
LENGTH: 1198
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-405

Query Match 37.3%; Score 47; DB 4; Length 1198;
Best Local Similarity 52.4%; Pred. No. 77;
Matches 11; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

OY 1 TDLQERGDNDISF---SGD 17
|||:|||||:|||
Db 400 TDLRPSIDNLRPFMAKXSGD 420

RESULT 5

US-08-341-843B-9
Sequence 9, Application US/08341843B
Patent No. 5872225
GENERAL INFORMATION:
APPLICANT: Lemmon, Vance

ATTORNEY/AGENT INFORMATION:
NAME: PARKER, DAVID L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: MAKE:002/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: na
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 348 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-669-435-2

Query Match 34.9% Score 44; DB 2; Length 348;
Best Local Similarity 53.3%; Pred. No. 53;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TDLOERGDNDISPFS 15
Db 320 TELQEKTDNMEPRS 334

RESULT 10
PCT-US94-14431A-2
Sequence 2, Application PC/TUS9414431A
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: IDEAL BILE ACID TRANSPORTER COMPOSITIONS
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P. O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: Patent Release #1.0, Version
SOFTWARE: #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/14431A
FILING DATE: 29-DEC-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/176,126
FILING DATE: 29-DEC-1993
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: PARKER, DAVID L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: MAKE005P--
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924(1) GENERAL INFORMATION:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 348 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-14431A-2

Query Match 34.9% Score 44; DB 5; Length 348;
Best Local Similarity 53.3%; Pred. No. 53;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TDLOERGDNDISPFS 15
Db 320 TELQEKTDNMEPRS 334

RESULT 11
US-08-278-091-6
Sequence 6, Application US/08278091
Patent No. 5506139
GENERAL INFORMATION:
APPLICANT: LOOSMORE, Sheena M
APPLICANT: YANG, Yan-Ping
APPLICANT: CHONG, Pele
APPLICANT: OOMEN, Raymond P.
APPLICANT: KLEIN, Michel H.
TITLE OF INVENTION: Analog of Haemophilus Hln47 Protein with
REDUCED PROTEASE ACTIVITY
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Slim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/278,091
FILING DATE: 21-JUL-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-371
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 475 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-278-091-6

Query Match 34.9% Score 44; DB 1; Length 475;
Best Local Similarity 57.1%; Pred. No. 76;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 10 DISFGDGOPEKD 23
Db 79 DQSPFCODSFPON 92

RESULT 12
US-08-483-859-6
Sequence 6, Application US/08483859
Patent No. 556436
GENERAL INFORMATION:
APPLICANT: LOOSMORE, Sheena M.
APPLICANT: YANG, Yan-Ping
APPLICANT: CHONG, Pele
APPLICANT: OOMEN, Raymond P.
APPLICANT: KLEIN, Michel H.
TITLE OF INVENTION: Analog of Haemophilus Hln47 Protein with
REDUCED PROTEASE ACTIVITY
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Slim & McBurney

STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,859
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/296,149
FILING DATE: 26-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/278,091
FILING DATE: 21-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-495 MIS.vg
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 475 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-483-859-6

Query Match 34.9%; Score 44; DB 1; Length 475;
Best Local Similarity 57.1%; Pred. No. 76;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 10 DISFSGDGPFRD 23
DB 79 DDSFPCDGSFON 92

RESULT 13
US-08-472-173-6
Sequence 6, Application US/08472173
Patent No. 565353
GENERAL INFORMATION:
APPLICANT: LOOSMORE, Sheena M
APPLICANT: YANG, Yan-Ping
APPLICANT: CHONG, Pele
APPLICANT: COHEN, Raymond P.
APPLICANT: KLEIN, Michel H.
TITLE OF INVENTION: Analog of Haemophilus Hinf7 Protein with
TITLE OF INVENTION: Reduced Protease Activity
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,173
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/296,149
FILING DATE: 26-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/278,091
FILING DATE: 21-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-493 MIS.vg
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 475 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-472-173-6

Query Match 34.9%; Score 44; DB 1; Length 475;
Best Local Similarity 57.1%; Pred. No. 76;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 10 DISFSGDGPFRD 23
DB 79 DDSFPCDGSFON 92

RESULT 14
US-08-350-741-2
Sequence 2, Application US/08350741
Patent No. 5804194
GENERAL INFORMATION:
APPLICANT: DOUGAN G.
APPLICANT: CHARLES I.G.
APPLICANT: HORMACHE C.E.
APPLICANT: JOHNSON K.S.
APPLICANT: CHATFIELD S.N.
TITLE OF INVENTION: LIVE VACCINES
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON and VANDERHAYE PC
STREET: 8th FLOOR, 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/350,741
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/952,737
FILING DATE: 09-MAY-1994
APPLICATION NUMBER: US 07/952,737
FILING DATE: 30-NOV-1992
APPLICATION NUMBER: GB 9007194.5
FILING DATE: 30-MAR-1990
APPLICATION NUMBER: PCT/GB91/00484
FILING DATE: 28-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: WILSON, MARY J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 117-158
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000

us-09-641-034-47.rai

Page 7

```
QY 10 DISFSGDGQPFKD 23
    | | | | | | | :
Db 79 DDSFQCQDGSPEQN 92
```

Query Match	34.98;	Score 44;	DB 1;	Length 475;
Best Local Similarity	57.18;	Pred. No. 76;		
Matches	8;	Conservative	4;	Indels 0;
				Gaps 0;

QY	10	DISPFGSGOPFKD	23
Db	79	DDSPFCODGSPFQD	92

Sequence 6, Application US/08487167
Patent No. 5869302
GENERAL INFORMATION:
APPLICANT: LOOSMORE, Sheena M.
APPLICANT: YANG, Yan-Ping
APPLICANT: CHONG, Pele
APPLICANT: DOMEN, Raymond P.
APPLICANT: KLEIN, Michel H.
TITLE OF INVENTION: Analog of Haemophilus Hinf7 Protein with
TITLE OF INVENTION: Reduced Protease Activity
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:

STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1G7

```

1 COMPUTER READABLE FORM:
2 MEDIUM TYPE: Floppy disk
3 COMPUTER: IBM PC compatible
4 OPERATING SYSTEM: PC-DOS/MS-DOS
5 SOFTWARE: Patentin Release #1.0, Version #1.25
6 CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/487,167
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/296,149
FILING DATE: 26-AUG-1994
PRIORITY INFORMATION:

FROM APPLICATION DATA:
APPLICATION NUMBER: US 08/278,091
FILING DATE: 21-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I.

REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-508 MIS:V9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE INFORMATION:

```

SEQUENCE CHARACTERISTICS:
LENGTH: 475 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-487-167-6

```

Query Match	34.98;	Score 44;	DB 2;	Length 475;
Best Local Similarity	57.18;	Pred. No. 76;		
Matches	8;	Conservative	2;	Mismatches 4;
				Indels 0;
				Gaps 0;

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 13:18:07 ; Search time 10 Seconds
(without alignments)
45.721 Million cell updates/sec

Title: US-09-641-034-47

Perfect score: 126

Sequence: 1 TDLQERGDNDISFSGDGPFDK 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 120991 seqs, 19878514 residues

Total number of hits satisfying chosen parameters: 120991

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications, RA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	126	100.0	23	9	US-09-812-485A-49
2	126	100.0	97	9	US-09-812-485A-1
3	126	100.0	525	10	US-09-814-550-2
4	101	80.2	40	9	US-09-812-485A-23
5	86	68.3	19	10	US-09-814-550-6
6	84	66.7	15	9	US-09-812-485A-48
7	84	66.7	33	9	US-09-812-485A-47
8	82	65.1	15	9	US-09-812-485A-30
9	75	59.5	15	9	US-09-812-485A-46
10	73.5	58.3	40	9	US-09-812-485A-26
11	72.5	57.5	30	9	US-09-812-485A-25
12	72.5	57.5	35	9	US-09-812-485A-42
13	66	52.4	12	9	US-09-812-485A-27
14	57	45.2	40	9	US-09-812-485A-32
15	57	45.2	45	9	US-09-812-485A-27
16	53	42.1	1458	12	US-10-054-691-2
17	52	41.3	19	10	US-09-814-550-5
18	52	41.3	301	10	US-09-905-176-23
19	49	38.9	9	10	US-09-814-550-8

20	49	38.9	326	9	US-09-738-626-4959	Sequence 4959, App
21	47	37.3	751	10	US-09-815-242-5832	Sequence 5832, App
22	47	37.3	825	10	US-09-815-242-12963	Sequence 12963, App
23	46.5	36.9	1137	12	US-10-100-912-4	Sequence 437, App
24	46.5	36.9	1141	9	US-09-978-295A-437	Sequence 437, App
25	46.5	36.9	1141	9	US-09-978-697-437	Sequence 437, App
26	46.5	36.9	1141	9	US-09-978-192A-437	Sequence 437, App
27	46.5	36.9	1141	9	US-09-999-882A-437	Sequence 437, App
28	46.5	36.9	1141	9	US-09-978-189-437	Sequence 437, App
29	46.5	36.9	1141	9	US-10-174-590-120	Sequence 120, App
30	46.5	36.9	1141	12	US-10-052-586-120	Sequence 120, App
31	46.5	36.9	1141	10	US-09-815-242-11959	Sequence 11959, App
32	45.5	36.1	1260	9	US-10-024-918-27	Sequence 27, App
33	45.5	36.1	9	10	US-09-814-550-7	Sequence 7, App
34	44	34.9	354	10	US-09-825-414-70	Sequence 70, App
35	43.5	34.1	1209	9	US-10-108-605-349	Sequence 349, App
36	43	34.1	2783	10	US-09-816-669A-14	Sequence 14, App
37	43	33.7	216	10	US-09-925-299-841	Sequence 841, App
38	42.5	33.7	545	10	US-09-878-242-1	Sequence 1, App
39	42.5	33.7	561	10	US-09-821-687-4	Sequence 4, App
40	42.5	33.7	579	10	US-09-925-300-1415	Sequence 1415, App
41	42.5	33.7	15	9	US-09-812-485A-45	Sequence 45, App
42	42	33.3	47	9	US-09-812-485A-2	Sequence 2, App
43	42	33.3	333	10	US-09-825-301-952	Sequence 952, App
44	42	33.3	639	10	US-09-815-242-5350	Sequence 5350, App
45	41.5	32.9	10			

ALIGNMENTS

RESULT 1
US-09-812-485A-49
Sequence 49, Application US/09812485A
Publication No. US20020197267A1
GENERAL INFORMATION:
APPLICANT: Kumagai, Yoshinari
APPLICANT: Blacher, Russel
TITLE OF INVENTION: Peptides and Methods of Treating Skeletal Diseases
FILE REFERENCE: BEAR-006CIP
CURRENT APPLICATION NUMBER: US/09/812,485A
CURRENT FILING DATE: 2001-03-19
PRIOR APPLICATION NUMBER: 09/641,034
PRIOR FILING DATE: 2000-08-16
NUMBER OF SEQ ID NOS: 50
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 49
LENGTH: 23
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: D-00006 peptide
NAME/KEY: AMIDATION
LOCATION: 15
US-09-812-485A-49

Query Match 100.0%; Score 126; DB 9; Length 23;
Best Local Similarity 100.0%; Pred No. 5.2e-13;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDLQERGDNDISFSGDGPFDK 23
DB 1 TDLQERGDNDISFSGDGPFDK 23
|||||
US-09-812-485A-1
; Sequence 1, Application US/09812485A
; Publication No. US20020197267A1
; GENERAL INFORMATION:
; APPLICANT: Kumagai, Yoshinari

APPLICANT: Blacher, Russel
APPLICANT: Yoneda, Toshiyuki
TITLE OF INVENTION: Integrin Binding Motif Containing
TITLE OF INVENTION: Peptides and Methods of Treating Skeletal Diseases
FILE REFERENCE: BEAR-006CIP
CURRENT FILING DATE: 2001-03-19
PRIOR APPLICATION NUMBER: 09/641,034
PRIOR FILING DATE: 2000-08-16
NUMBER OF SEQ ID NOS: 50
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 97
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Peptidic compound
US-09-812-485A-1

Query Match 100.0%; Score 126; DB 9; Length 97;
Best Local Similarity 100.0%; Pred. No. 2.6e-12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDQERGDNDISFSGDGPFRD 23
DB 43 TDQERGDNDISFSGDGPFRD 65

RESULT 3
US-09-814-550-2
Sequence 2, Application US/09814550
Patent No. US20020102641A1
GENERAL INFORMATION:
APPLICANT: Schlavy, Susan
APPLICANT: Madden, Stephen
APPLICANT: Manavalan, Parthasarathy
APPLICANT: Levine, Michael
APPLICANT: Jan de Beur, Suzanne
TITLE OF INVENTION: ONCOGENIC OSTEOMALACIA-RELATED GENE 1
FILE REFERENCE: 501405
CURRENT APPLICATION NUMBER: US/09/814,550
CURRENT FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: US 60/191,786
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: US 60/241,598
PRIOR FILING DATE: 2000-10-19
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 525
TYPE: PRT
ORGANISM: Homo sapiens
US-09-814-550-2

Query Match 100.0%; Score 126; DB 10; Length 525;
Best Local Similarity 100.0%; Pred. No. 1.8e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDQERGDNDISFSGDGPFRD 23
DB 242 TDQERGDNDISFSGDGPFRD 264

RESULT 4
US-09-812-485A-23
Sequence 23, Application US/09812485A
Publication No. US20020197267A1
GENERAL INFORMATION:
APPLICANT: Kumagai, Yoshinari
APPLICANT: Blacher, Russel
APPLICANT: Yoneda, Toshiyuki
TITLE OF INVENTION: Integrin Binding Motif Containing
TITLE OF INVENTION: Peptides and Methods of Treating Skeletal Diseases

FILE REFERENCE: BEAR-006CIP
CURRENT APPLICATION NUMBER: US/09/812,485A
CURRENT FILING DATE: 2001-03-19
PRIOR APPLICATION NUMBER: 09/641,034
PRIOR FILING DATE: 2000-08-16
NUMBER OF SEQ ID NOS: 50
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 23
LENGTH: 40
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Peptidic compound
US-09-812-485A-23

Query Match 80.2%; Score 101; DB 9; Length 40;
Best Local Similarity 100.0%; Pred. No. 5.6e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 RGDNDISFSGDGPFRD 23
DB 1 RGDNDISFSGDGPFRD 18

RESULT 5
US-09-814-550-6
Sequence 6, Application US/09814550
Patent No. US20020102641A1
GENERAL INFORMATION:
APPLICANT: Schlavy, Susan
APPLICANT: Madden, Stephen
APPLICANT: Manavalan, Parthasarathy
APPLICANT: Levine, Michael
APPLICANT: Jan de Beur, Suzanne
TITLE OF INVENTION: ONCOGENIC OSTEOMALACIA-RELATED GENE 1
FILE REFERENCE: 501405
CURRENT APPLICATION NUMBER: US/09/814,550
CURRENT FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: US 60/191,786
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: US 60/241,598
PRIOR FILING DATE: 2000-10-19
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.0
SEQ ID NO 6
LENGTH: 19
TYPE: PRT
ORGANISM: Homo sapiens
US-09-814-550-6

Query Match 68.3%; Score 86; DB 10; Length 19;
Best Local Similarity 100.0%; Pred. No. 4.4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 DLQERGDNDISFSGD 17
DB 4 DLQERGDNDISFSGD 19

RESULT 6
US-09-812-485A-48
Sequence 48, Application US/09812485A
Publication No. US20020197267A1
GENERAL INFORMATION:
APPLICANT: Kumagai, Yoshinari
APPLICANT: Blacher, Russel
APPLICANT: Yoneda, Toshiyuki
TITLE OF INVENTION: Integrin Binding Motif Containing
TITLE OF INVENTION: Peptides and Methods of Treating Skeletal Diseases
FILE REFERENCE: BEAR-006CIP
CURRENT APPLICATION NUMBER: US/09/812,485A
CURRENT FILING DATE: 2001-03-19
PRIOR APPLICATION NUMBER: 09/641,034

;; PRIOR FILING DATE: 2000-08-16
;; NUMBER OF SEQ ID NOS: 50
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 48
;; LENGTH: 15
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: D-00005 peptide
;; NAME/KEY: AMIDATION
;; LOCATION: 15
US-09-812-485A-48

Query Match
Best Local Similarity 66.7%; Score 84; DB 9; Length 15;
Matches 15; Conservative 100.0%; Pred. No. 6,7e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 NDSPFSGDGPDKD 23
DB 1 NDSPFSGDGPDKD 15

RESULT 7
US-09-812-485A-30
;; Sequence 30, Application US/09812485A
;; Publication No. US20020197267A1
;; GENERAL INFORMATION:
;; APPLICANT: Kumagai, Yoshinari
;; APPLICANT: Blacher, Russel
;; APPLICANT: Yoneda, Toshiyuki
;; TITLE OF INVENTION: Integrin Binding Motif Containing
;; FILE REFERENCE: BEAR-006CIP
;; CURRENT APPLICATION NUMBER: US/09/812,485A
;; PRIOR FILING DATE: 2001-03-19
;; PRIOR APPLICATION NUMBER: 09/641,034
;; NUMBER OF SEQ ID NOS: 50
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 30
;; LENGTH: 33
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: peptidic compound
US-09-812-485A-30

Query Match
Best Local Similarity 66.7%; Score 84; DB 9; Length 33;
Matches 15; Conservative 100.0%; Pred. No. 1,6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 NDSPFSGDGPDKD 23
DB 1 NDSPFSGDGPDKD 15

RESULT 8
US-09-812-485A-47
;; Sequence 47, Application US/09812485A
;; Publication No. US20020197267A1
;; GENERAL INFORMATION:
;; APPLICANT: Kumagai, Yoshinari
;; APPLICANT: Blacher, Russel
;; APPLICANT: Yoneda, Toshiyuki
;; TITLE OF INVENTION: Integrin Binding Motif Containing
;; FILE REFERENCE: BEAR-006CIP
;; CURRENT APPLICATION NUMBER: US/09/812,485A
;; PRIOR FILING DATE: 2001-03-19
;; PRIOR APPLICATION NUMBER: 09/641,034
;; NUMBER OF SEQ ID NOS: 50
;; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 47
;; LENGTH: 15
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: D-00004 peptide
;; NAME/KEY: AMIDATION
;; LOCATION: 15
US-09-812-485A-47

Query Match
Best Local Similarity 65.1%; Score 82; DB 9; Length 15;
Matches 15; Conservative 100.0%; Pred. No. 1,3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 ERGDNDISPSGDCQ 19
DB 1 ERGDNDISPSGDCQ 15

RESULT 9
US-09-812-485A-46
;; Sequence 46, Application US/09812485A
;; Publication No. US20020197267A1
;; GENERAL INFORMATION:
;; APPLICANT: Kumagai, Yoshinari
;; APPLICANT: Blacher, Russel
;; APPLICANT: Yoneda, Toshiyuki
;; TITLE OF INVENTION: Integrin Binding Motif Containing
;; FILE REFERENCE: BEAR-006CIP
;; CURRENT APPLICATION NUMBER: US/09/812,485A
;; PRIOR FILING DATE: 2001-03-19
;; PRIOR APPLICATION NUMBER: 09/641,034
;; NUMBER OF SEQ ID NOS: 50
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 46
;; LENGTH: 15
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: D-00003 peptide
;; NAME/KEY: AMIDATION
;; LOCATION: 15
US-09-812-485A-46

Query Match
Best Local Similarity 59.5%; Score 75; DB 9; Length 15;
Matches 14; Conservative 100.0%; Pred. No. 1,5e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDQERGDNDISPF 14
DB 2 TDQERGDNDISPF 15

RESULT 10
US-09-812-485A-24
;; Sequence 24, Application US/09812485A
;; Publication No. US20020197267A1
;; GENERAL INFORMATION:
;; APPLICANT: Kumagai, Yoshinari
;; APPLICANT: Blacher, Russel
;; APPLICANT: Yoneda, Toshiyuki
;; TITLE OF INVENTION: Integrin Binding Motif Containing
;; FILE REFERENCE: BEAR-006CIP
;; CURRENT APPLICATION NUMBER: US/09/812,485A
;; PRIOR FILING DATE: 2001-03-19
;; PRIOR APPLICATION NUMBER: 09/641,034
;; NUMBER OF SEQ ID NOS: 50
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 24

LENGTH: 40
TYPE: PRF
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: peptidic compound
US-09-812-485A-24

Query Match 58.3%; Score 73.5; DB 9; Length 40;
Best Local Similarity 83.3%; Pred. No. 7.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

OY 6 RGDNDISFGSGDGPFD 23
DB 4 RGD---SPFGSGDGPFD 18

RESULT 11
US-09-812-485A-26
Sequence 26, Application US/09812485A
Publication No. US20020197267A1
GENERAL INFORMATION:
APPLICANT: Kumagai, Yoshinari
APPLICANT: Blacher, Russel
APPLICANT: Yoneda, Toshiyuki
TITLE OF INVENTION: Integrin Binding Motif Containing
Peptides and Methods of Treating Skeletal Diseases
FILE REFERENCE: BEAR-006CIP
CURRENT APPLICATION NUMBER: US/09/812,485A
CURRENT FILING DATE: 2001-03-19
PRIOR APPLICATION NUMBER: 09/641,034
PRIOR FILING DATE: 2000-08-16
NUMBER OF SEQ ID NOS: 50
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 26
LENGTH: 30
TYPE: PRF
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: peptidic compound
US-09-812-485A-26

Query Match 57.5%; Score 72.5; DB 9; Length 30;
Best Local Similarity 83.3%; Pred. No. 7.8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

OY 9 NDISPFS---GDGPFKD 23
DB 1 NDISPFSGDGDPFD 18

RESULT 12
US-09-812-485A-25
Sequence 25, Application US/09812485A
Publication No. US20020197267A1
GENERAL INFORMATION:
APPLICANT: Kumagai, Yoshinari
APPLICANT: Blacher, Russel
APPLICANT: Yoneda, Toshiyuki
TITLE OF INVENTION: Integrin Binding Motif Containing
Peptides and Methods of Treating Skeletal Diseases
FILE REFERENCE: BEAR-006CIP
CURRENT APPLICATION NUMBER: US/09/812,485A
CURRENT FILING DATE: 2001-03-19
PRIOR APPLICATION NUMBER: 09/641,034
PRIOR FILING DATE: 2000-08-16
NUMBER OF SEQ ID NOS: 50
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 25
LENGTH: 35
TYPE: PRF
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: peptidic compound

US-09-812-485A-25

Query Match 57.5%; Score 72.5; DB 9; Length 35;
Best Local Similarity 83.3%; Pred. No. 9.3e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

OY 9 NDISPFS---SGDGPFD 23
DB 1 NDISPFRDSDGDPFD 18

RESULT 13
US-09-812-485A-42
Sequence 42, Application US/09812485A
Publication No. US20020197267A1
GENERAL INFORMATION:
APPLICANT: Kumagai, Yoshinari
APPLICANT: Blacher, Russel
APPLICANT: Yoneda, Toshiyuki
TITLE OF INVENTION: Integrin Binding Motif Containing
Peptides and Methods of Treating Skeletal Diseases
FILE REFERENCE: BEAR-006CIP
CURRENT APPLICATION NUMBER: US/09/812,485A
CURRENT FILING DATE: 2001-03-19
PRIOR APPLICATION NUMBER: 09/641,034
PRIOR FILING DATE: 2000-08-16
NUMBER OF SEQ ID NOS: 50
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 42
LENGTH: 12
TYPE: PRF
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: calcium binding motif
US-09-812-485A-42

Query Match 52.4%; Score 66; DB 9; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 DNDISFGSDGQ 19
DB 1 DNDISFGSDGQ 12

RESULT 14
US-09-812-485A-32
Sequence 32, Application US/09812485A
Publication No. US20020197267A1
GENERAL INFORMATION:
APPLICANT: Kumagai, Yoshinari
APPLICANT: Blacher, Russel
APPLICANT: Yoneda, Toshiyuki
TITLE OF INVENTION: Integrin Binding Motif Containing
Peptides and Methods of Treating Skeletal Diseases
FILE REFERENCE: BEAR-006CIP
CURRENT APPLICATION NUMBER: US/09/812,485A
CURRENT FILING DATE: 2001-03-19
PRIOR APPLICATION NUMBER: 09/641,034
PRIOR FILING DATE: 2000-08-16
NUMBER OF SEQ ID NOS: 50
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 32
LENGTH: 40
TYPE: PRF
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: peptidic compound
US-09-812-485A-32

Query Match 45.2%; Score 57; DB 9; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.023;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 14 FSGDGOPEKD 23
 Db 1 FSGDGOPEKD 10

RESULT 15

US-09-812-485A-27
 ; Sequence 27, Application US/09812485A
 ; Publication No. US20020197267A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumagai, Yoshiaki
 ; APPLICANT: Blacher, Russel
 ; APPLICANT: Yoneda, Yoshiyuki
 ; TITLE OF INVENTION: Integrin Binding Motif Containing
 ; TITLE OF INVENTION: Peptides and Methods of Treating Skeletal Diseases
 ; FILE REFERENCE: BEAR-006CIP
 ; CURRENT APPLICATION NUMBER: US/09/812,485A
 ; CURRENT FILING DATE: 2001-03-19
 ; PRIOR APPLICATION NUMBER: 09/641,034
 ; PRIOR FILING DATE: 2000-08-16
 ; NUMBER OF SEQ ID NOS: 50
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 27
 ; LENGTH: 45
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: peptidic compound
 US-09-812-485A-27

Query Match 45.2%; Score 57; DB 9; Length 45;
 Best Local Similarity 100.0%; Pred. No. 0.027;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 14 FSGDGOPEKD 23
 Db 1 FSGDGOPEKD 10

Search completed: January 17, 2003, 13:22:03
 Job time : 10 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 13:17:52 ; Search time 15 Seconds

(without alignments)
112,414 Million cell updates/sec

Title: US-09-641-034-47

Perfect score: 126
Sequence: 1 TDQERGDNDISPFSGDGPFKD 23Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 273288 seqs, 73313625 residues

Total number of hits satisfying chosen parameters: 273288

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA New:*

- 1: /cgn2_6/ptodata/2/paa/PCF_NEW_COMB.pep:*
- 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
- 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
- 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
- 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
- 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
- 7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	37.3	815	6	US-10-092-411A-3614
2	46.5	36.9	1141	6	US-10-145-087A-437
3	46.5	36.9	1141	6	US-10-143-031A-437
4	46.5	36.9	1141	6	US-10-145-092A-437
5	46.5	36.9	1141	6	US-10-162-522A-437
6	46.5	36.9	1141	6	US-10-165-038A-437
7	46.5	36.9	1141	6	US-10-165-353-437
8	46.5	36.9	1141	6	US-10-170-481A-437
9	46.5	36.9	1141	6	US-10-172-039A-437
10	46.5	36.9	1141	6	US-10-145-016A-437
11	46.5	36.9	1141	6	US-10-145-088A-437
12	46.5	36.9	1141	6	US-10-145-129A-437
13	46.5	36.9	1141	6	US-10-125-823A-120
14	46.5	36.9	1141	6	US-10-165-353A-437
15	46.5	36.9	1141	6	US-10-205-892-120
16	46.5	36.9	1141	6	US-10-174-575-120
17	46.5	36.9	1141	6	US-10-174-575A-120
18	46.5	36.9	1141	6	US-10-162-521A-437
19	46.5	36.9	1141	6	US-10-167-600-437
20	46.5	36.9	1141	6	US-10-187-755-120
21	46.5	36.9	1141	6	US-10-013-924A-437
22	46.5	36.9	1141	6	US-10-013-928A-437
23	46.5	36.9	1141	6	US-10-167-600A-437
24	46.5	36.9	1141	6	US-10-187-749-120
25	46.5	36.9	1165	5	US-09-724-676-75168
26	46.5	36.9	1165	5	US-09-724-676-75168

27	46.5	36.9	1165	5	US-09-724-676A-75168	Sequence 75168, A
28	46.5	36.9	1172	5	US-09-724-676-75166	Sequence 75166, A
29	46.5	36.9	1172	5	US-09-724-676A-75166	Sequence 75166, A
30	46.5	36.9	1200	5	US-09-724-676-75170	Sequence 75170, A
31	46.5	36.9	1200	5	US-09-724-676A-75170	Sequence 75170, A
32	46.5	36.9	1203	5	US-09-724-676-75169	Sequence 75169, A
33	46.5	36.9	1210	5	US-09-724-676A-75169	Sequence 75169, A
34	46.5	36.9	1210	5	US-09-724-676-75167	Sequence 75167, A
35	46.5	36.9	1238	5	US-09-724-676-75171	Sequence 75171, A
36	46.5	36.9	1238	5	US-09-724-676A-75171	Sequence 75171, A
37	46.5	36.9	1238	5	US-09-724-676-75171	Sequence 75171, A
38	46	36.5	648	1	PCT-US02-32727-15272	Sequence 15272, A
39	46	36.5	648	6	US-10-057-498-15272	Sequence 15272, A
40	43	34.1	2476	5	US-09-724-676-91032	Sequence 91032, A
41	43	34.1	2476	5	US-09-724-676A-91032	Sequence 91032, A
42	43	34.1	2548	5	US-09-724-676-91030	Sequence 91030, A
43	43	34.1	2548	5	US-09-724-676A-91030	Sequence 91030, A
44	43	34.1	2826	5	US-09-724-676-91028	Sequence 91028, A
45	43	34.1	2826	5	US-09-724-676A-91028	Sequence 91028, A

ALIGNMENTS

RESULT 1
US-10-092-411A-3614
Sequence 3614, Application US/10092411A
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 032796-101
CURRENT APPLICATION NUMBER: US/10/092.411A
CURRENT FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: US 09/134,001
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5676
SEQ ID NO 3614
LENGTH: 815
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-10-092-411A-3614

Query Match 37.3%; Score 47; DB 6; Length 815;
Best Local Similarity 58.8%; Pred. No. 78;
Matches 10; Conservative 7; Indels 0; Gaps 0;

QY 7 GDNDISPFSGDGPFKD 23
DB 481 GDEDLPVSGDGPVKD 497
RESULT 2
US-10-145-087A-437
Sequence 437, Application US/10145087A
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avl
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Ealon, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gottlieb, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.

```

: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kijavlin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James.
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2630PIC47
: CURRENT APPLICATION NUMBER: US/10/145,087A
: PRIOR APPLICATION NUMBER: 2001-10-18
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 624
: SEQ ID NO 437
: LENGTH: 1141
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-145-087A-437
```

```

Query Match          36.9%; Score 46.5; DB 6; Length 1141;
Best Local Similarity 47.8%; Pred. No. 1.3e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;
```

OY 2 DLQERGDNDI---SPFSGDGQPF 21

DB 394 DLNODGFPDIAGAPFDCGKVF 416

RESULT 3
US-10-143-031A-437

: Sequence 437, Application US/10143031A

: GENERAL INFORMATION:

```

: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
```

```

: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kijavlin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James.
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2630PIC39
: CURRENT APPLICATION NUMBER: US/10/143,031A
: PRIOR APPLICATION NUMBER: 2002-10-10
: PRIOR FILING DATE: 2002-10-10
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 624
: SEQ ID NO 437
: LENGTH: 1141
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-143-031A-437
```

```

Query Match          36.9%; Score 46.5; DB 6; Length 1141;
Best Local Similarity 47.8%; Pred. No. 1.3e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;
```

OY 2 DLQERGDNDI---SPFSGDGQPF 21

DB 394 DLNODGFPDIAGAPFDCGKVF 416

RESULT 4
US-10-145-092A-437

: Sequence 437, Application US/10145092A

: GENERAL INFORMATION:

```

: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
```



```

: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kljavin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James.
: APPLICANT: Paonli, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2630PIC29
: CURRENT APPLICATION NUMBER: US/10/165,038A
: PRIOR FILING DATE: 2002-10-10
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 624
: SEQ ID NO 437
: LENGTH: 1141
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-165-038A-437

Query Match          36.9% Score 46.5% DB 6: Length 1141:
Best Local Similarity 47.8% Pred. No. 1.3e+02:
Matches 11: Conservative 3: Mismatches 6: Indels 3: Gaps 1:

OY 2 DLOERGDNDI---SPFGSGQPF 21
||:| |||: ||| |||: |
Db 394 DLNODGFPDIAGVAFDGDGKVF 416

RESULT 7
US-10-165-353-437
: Sequence 437, Application US/10165353
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
```

```

: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kljavin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James.
: APPLICANT: Paonli, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2630PIC40
: CURRENT APPLICATION NUMBER: US/10/165,353
: PRIOR FILING DATE: 2002-10-10
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 624
: SEQ ID NO 437
: LENGTH: 1141
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-165-353-437

Query Match          36.9% Score 46.5% DB 6: Length 1141:
Best Local Similarity 47.8% Pred. No. 1.3e+02:
Matches 11: Conservative 3: Mismatches 6: Indels 3: Gaps 1:

OY 2 DLOERGDNDI---SPFGSGQPF 21
||:| |||: ||| |||: |
Db 394 DLNODGFPDIAGVAFDGDGKVF 416

RESULT 8
US-10-170-481A-437
: Sequence 437, Application US/10170481A
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
```

```

: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kljavin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James.
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2630P1C53
: CURRENT APPLICATION NUMBER: US/10/170,481A
: PRIOR FILING DATE: 2002-10-10
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 624
: SEQ ID NO 437
: LENGTH: 1141
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-170-481A-437

Query Match          36.9%; Score 46.5; DB 6; Length 1141;
Best Local Similarity 47.8%; Pred. No. 1.3e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

QY      2 DLQERGDNDI---SPFSGDQGF 21
      11 : 1 11 : 11 111 : 1
Db      394 DLNDGFPDIAVGAPFDGDGKVF 416

RESULT 9
US-10-172-039A-437
: Sequence 437, Application US/10172039A
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
```

```

: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kljavin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James.
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2630P1C30
: CURRENT APPLICATION NUMBER: US/10/172,039A
: PRIOR FILING DATE: 2002-10-10
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 624
: SEQ ID NO 437
: LENGTH: 1141
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-172-039A-437

Query Match          36.9%; Score 46.5; DB 6; Length 1141;
Best Local Similarity 47.8%; Pred. No. 1.3e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

QY      2 DLQERGDNDI---SPFSGDQGF 21
      11 : 1 11 : 11 111 : 1
Db      394 DLNDGFPDIAVGAPFDGDGKVF 416

RESULT 10
US-10-145-016A-437
: Sequence 437, Application US/10145016A
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
```

```

: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kijavito, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James.
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2630P1C49
: CURRENT APPLICATION NUMBER: US/10/145.016A
: PRIOR FILING DATE: 2001-10-18
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 624
: SEQ ID NO 437
: LENGTH: 1141
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-145-016A-437

Query Match          36.9%; Score 46.5; DB 6; Length 1141;
Best Local Similarity 47.8%; Pred. No. 1.3e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

OY      2 DLORGNDR---SPFGDQGF 21
Db      394 DLNODGFPIAVGAPFDGDKVF 416

RESULT 11
: Sequence 437, Application US/10145088A
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gertlisen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.

```

```

: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kijavito, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James.
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: P2630P1C49
: CURRENT APPLICATION NUMBER: US/10/145.088A
: PRIOR FILING DATE: 2002-10-10
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 624
: SEQ ID NO 437
: LENGTH: 1141
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-145-088A-437

Query Match          36.9%; Score 46.5; DB 6; Length 1141;
Best Local Similarity 47.8%; Pred. No. 1.3e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

OY      2 DLORGNDR---SPFGDQGF 21
Db      394 DLNODGFPIAVGAPFDGDKVF 416

RESULT 12
: Sequence 437, Application US/10145129A
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gertlisen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.

```

```

; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Acids and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC51
; CURRENT APPLICATION NUMBER: US/10/145,129A
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 624
; SEQ ID NO 437
; LENGTH: 1141
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-145-129A-437

Query Match      36.9%; Score 46.5; DB 6; Length 1141;
Best Local Similarity 47.8%; Pred. No. 1.3e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

QY      2 DLQERGDNDI---SPFSGDCQPF 21
      11 : 1 11 : 11 111 : 1
Db      394 DLNODGEPDIAVGAPFDGDKVF 416

RESULT 13
US-10-125-923A-120
; Sequence 120, Application US/10125923A
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME

```

```

; FILE REFERENCE: P3430RIC79
; CURRENT APPLICATION NUMBER: US/10/125,923A
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 120
; LENGTH: 1141
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-125-923A-120

Query Match      36.9%; Score 46.5; DB 6; Length 1141;
Best Local Similarity 47.8%; Pred. No. 1.3e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

QY      2 DLQERGDNDI---SPFSGDCQPF 21
      11 : 1 11 : 11 111 : 1
Db      394 DLNODGEPDIAVGAPFDGDKVF 416

RESULT 14
US-10-165-353A-437
; Sequence 437, Application US/10165353A
; GENERAL INFORMATION:
; APPLICANT: Asukenzai, AVI
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same

```

```
FILE REFERENCE: P2630P1C40
CURRENT APPLICATION NUMBER: US/10/165.353A
CURRENT FILING DATE: 2002-10-10
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
Remaining Prior Application data removed - See File Wrapper or PALM.
SEQU ID NO 437
LENGTH: 1141
TYPE: PRT
ORGANISM: Homo sapiens
US-10-165-353A-437
```

```
Query Match 36.9%; Score 46.5; DB 6; Length 1141;
Best Local Similarity 47.8%; Pred. No. 1.3e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;
```

```
OY 2 DLQERGDNDI---SPFSGDGP 21
DB 394 DLNODGFPDIAVGAPFDGDKVF 416
```

```
RESULT 15
US-10-205-892-120
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C517
CURRENT APPLICATION NUMBER: US/10/205.892
CURRENT FILING DATE: 2002-07-26
PRIOR APPLICATION NUMBER: 10/052586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
```

```
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
Remaining Prior Application data removed - See File Wrapper or PALM.
SEQU ID NO 120
LENGTH: 1141
TYPE: PRT
ORGANISM: Homo Sapien
US-10-205-892-120
```

```
Query Match 36.9%; Score 46.5; DB 6; Length 1141;
Best Local Similarity 47.8%; Pred. No. 1.3e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;
```

```
OY 2 DLQERGDNDI---SPFSGDGP 21
DB 394 DLNODGFPDIAVGAPFDGDKVF 416
```

```
Search completed: January 17, 2003, 13:21:47
Job time : 17 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 13:16:38 ; Search time 16 seconds

(without alignments)
138.193 Million cell updates/sec

Title: US-09-641-034-47

Sequence: 126
1 TDQERGDNDISPPSGDGPFRKD 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR73: *
2: PIR1: *
3: PIR2: *
4: PIR3: *
5: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	41.3	301	2	SI6070
2	52	41.3	320	2	JC5727
3	49.5	39.3	1259	2	S36126
4	48	38.1	277	2	AD3322
5	48	38.1	385	2	D72858
6	48	38.1	385	2	T41811
7	47	37.3	444	2	B83891
8	47	37.3	481	2	AH0410
9	47	37.3	825	2	B89940
10	46.5	36.9	1106	2	S38783
11	46.5	36.9	1135	2	161186
12	46.5	36.9	1137	2	JC5950
13	46	36.5	308	2	AD3154
14	46	36.5	365	2	E98133
15	45.5	36.1	347	2	A83170
16	45.5	36.1	1260	1	S05479
17	45	35.7	158	2	A11365
18	45	35.7	357	2	C84856
19	45	35.7	393	2	S59499
20	45	35.7	472	2	J00802
21	45	35.7	913	2	T38173
22	45	35.7	933	2	D82885
23	44.5	35.3	698	2	D65210
24	44.5	35.3	698	2	D91255
25	44.5	35.3	698	2	H86095
26	44	34.9	348	2	A49876
27	44	34.9	475	1	S15337
28	44	34.9	475	2	AC0528
29	44	34.9	683	2	D83511

30	43.5	34.5	284	2	C71507	probable L2 riboso
31	43.5	34.5	284	2	A81661	ribosomal protein
32	43.5	34.5	631	2	D38162	cobT protein - pse
33	43.5	34.5	1053	2	S44250	Integrin alpha-5 c
34	43.5	34.5	1115	2	T09403	Integrin alpha cha
35	43.5	34.5	1115	2	T09433	Integrin alpha cha
36	43	34.1	140	2	G84608	En/Spm-like transp
37	43	34.1	300	2	F90274	hypothetical prote
38	43	34.1	335	2	G72258	conserved hypotnet
39	43	34.1	629	2	T39285	probable transmemb
40	43	34.1	684	1	RNA6VS	transcription init
41	43	34.1	684	2	G97619	RNA polymerase sig
42	43	34.1	684	2	AF2842	RNA polymerase sig
43	43	34.1	762	2	E87592	hypothetical prote
44	43	34.1	779	2	G87573	xanthine dehydroge
45	43	34.1	891	2	E96590	hypothetical prote

ALIGNMENTS

RESULT 1
SI6070
porin - Rhodobacter capsulatus
C:Species: Rhodobacter capsulatus
C:Date: 21-Nov-1993 #sequence_revision 21-Jul-1995 #text_change 21-Jul-1995
C:Accession: SI6070
R:Schultz, E.; Kreusch, A.; Nestel, U.; Schulz, G.E.
Eur. J. Biochem. 199, 587-594, 1991
A:Title: Primary structure of porin from Rhodobacter capsulatus.
A:Reference number: SI6070; MUID:91330909; PMID:1651239
A:Accession: SI6070
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-301 <EUF>
A:Experimental source: strain 37b4
A>Note: the source is designated as Rhodobacter capsulatus

Query Match 41.3% Score 52; DB 2; Length 301;
Best Local Similarity 58.8% Pred. No. 3.1;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
OY 1 TDQERGDNDISPPSGD 17
DB 92 TDLDGRGNDIPYLTGD 108

RESULT 2
JC5727
porin precursor - Rhodobacter capsulatus
C:Species: Rhodobacter capsulatus
C:Date: 09-Dec-1997 #sequence_revision 23-Jan-1998 #text_change 07-May-1999
C:Accession: JC5727; PC4414
R:Trieschmann, M.D.A.; Patus, F.; Tadros, M.H.
Gene 183, 61-68, 1996
A:Title: Molecular characterization and organization of porin from Rhodobacter capsu.

A:Reference number: JC5727; MUID:91149280; PMID:8996088
A:Accession: JC5727

A:Molecule type: DNA

A:Residues: 1-320 <TRI>

A:Cross-references: GB:U57653

A:Experimental source: strain 37B4

A:Accession: PC4414

A:Molecule type: protein

A:Residues: 67-89;131-158;191-218;252-277 <TR2>

C:Comment: This protein is the mutant of porin from Rhodobacter capsulatus strain 37I

A:Gene: porCa

F:1-20/Domain: signal sequence #status predicted <SIG>

F:21-320/Product: porin #status predicted <MNT>

Query Match

41.3% Score 52; DB 2; Length 320;

Best Local Similarity 58.8%; Pred. No. 3.3;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
OY 1 TDJERGNDISFSGD 17
|||:|||||:|
DB 112 TDJDRGNDIPYLTGD 128

RESULT 3

S36126

neural cell adhesion molecule L1 - rat

M:Alternate names: nerve growth factor-inducible large external glycoprotein; NILE glyco

C:Species: Rattus norvegicus (Norway rat)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Aug-1999

R:Accession: S36126; S17655; A60917; A30326

R:Miura, M.; Kobayashi, M.; Asou, H.; Umemura, K.

A:Title: Molecular cloning of cDNA encoding the rat neural cell adhesion molecule L1. T

A:Reference number: S17655; MUID:91372414; PMID:1894011

A:Accession: S36126

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1259 <MID>

A:CROSS-references: EMBL:X59149

A:Accession: S17655

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1178, 1183-1259 <MID>

A:CROSS-references: EMBL:X59149; NID:956740; PIDN:CMAA1860.1; PID:956741

R:Prince, J.T.; Milona, N.; Stallcup, W.B.

J. Neurosci. 9, 1825-1834, 1989

A:Title: Characterization of a partial cDNA clone for the NILE glycoprotein and identifi

A:Reference number: A60917; MUID:89257627; PMID:2723751

A:Accession: A60917

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1159-1199, 'G', 1201-1235, 'K', 1237 <PRI>

A:Note: this paper appeared earlier, with printing errors, as reference A30326

R:Prince, J.T.; Milona, N.; Stallcup, W.B.

J. Neurosci. 9, 876-883, 1989

A:Title: Characterization of a partial cDNA clone for the NILE glycoprotein and identifi

A:Reference number: A30326; MUID:89177485; PMID:2466966

A:Contents: annotation

A:Note: this paper was reprinted as reference A60917 to correct the omission of several

C:Comment: This sequence of this surface-accessible glycoprotein differs at only two pos

C:Superfamily: neural cell adhesion molecule L1; fibronectin type III repeat homology; A

C:Keywords: cell adhesion; duplication; glycoprotein; membrane protein

F:531-592/Domains: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 39.3%; Score 49.5; DB 2; Length 1259;

Matches 12; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

OY 2 DLOERGNDISFSGDQ 19

|||||||:|

DB 558 DLOERGNDISFSGDQ 574

RESULT 4

AD3322

lysosome (BC 3.2.1.17) - Brucella melitensis (strain 16M)

C:Species: Brucella melitensis

C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 19-Apr-2002

C:Accession: AD3322

R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanova,

M.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Leless

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens

A:Reference number: AD3322; PMID:11756688

A:Accession: AD3322

A:Status: preliminary

A:Molecule type: DNA

A:Experimental source: isolate T3

A:Residues: 1-277 <KUR>
A:CROSS-references: GB:AEO08917; PIDN:AAL51743.1; PID:g17982481; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics: BME10562
A:Gene: BME10562
A:Map position: I
C:Keywords: glycosidase; hydrolase

Query Match

Best Local Similarity 38.1%; Score 48; DB 2; Length 277;

Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 6 RGDNDISFSGDOPFK 22

|||:|||||:|

DB 252 RGDNDISFSGDOPFK 268

RESULT 5

D72858

late expression factor 3 - Autographa californica nuclear polyhedrosis virus

C:Species: Autographa californica nuclear polyhedrosis virus, AcMNPV

A:Note: dsDNA virus

C:Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 24-Nov-1999

R:Accession: D72858; A40677

R:Avres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R.D.

A:Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis vir

A:Reference number: A72850; MUID:94303173; PMID:8030224

A:Accession: D72858

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-385 <AVR>

A:CROSS-references: GB:L22858; NID:9510708; PIDN:AAA66697.1; PID:9559136

R:Li, Y.; Passarelli, A.L.; Miller, L.K.

J. Virol. 67, 5260-5268, 1993

A:Title: Identification, sequence, and transcriptional mapping of Ief-3, a baculoviru

A:Reference number: A40677; MUID:93353600; PMID:8350397

A:Accession: A40677

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-385 <LAI>

A:CROSS-references: GB:L18873; NID:9349019; PIDN:AAA02964.1; PID:9349020

A:Note: authors translated the codon ACA for residue 92 as Tyr

C:Genetics: Ief3

C:Keywords: transcription regulation

Query Match

Best Local Similarity 38.1%; Score 48; DB 2; Length 385;

Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 8 DNDISFSGDOPFKD 23

|||:|||||:|

DB 61 DNDISFSGDOPFKD 76

RESULT 6

T41811

LEF-3 orf67 - Bombyx mori nuclear polyhedrosis virus (isolate T3)

C:Species: Bombyx mori nuclear polyhedrosis virus, BmNPV

A:Variety: isolate T3

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000

C:Accession: T41811

R:Gomi, S.; Maizumi, K.; Maeda, S.

J. Gen. Virol. 80, 1323-1337, 1999

A:Title: Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus.

A:Reference number: 22020; MUID:93281911; PMID:10355780

A:Accession: T41811

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-385 <KAM>

A:CROSS-references: EMBL:L33180; NID:93745835; PIDN:AAAC63740.1; PID:93745893

A:Experimental source: isolate T3

C:Genetics:
A:Note: lrf-3

Query Match 38.1%; Score 48; DB 2; Length 385;
Best Local Similarity 56.2%; Pred. No. 17;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 8 DNDISPFSGDGPFRD 23
DB 61 DNKIOEYXGDSQSFKD 76

RESULT 7

B83891
Intercellular alkaline serine proteinase aprX [Imported] - Bacillus halodurans (strain C
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: B83891

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: B83891

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-444 <STO>
A:Cross-references: GB:BA000004; NID:g10174345; PIDN:BA05649.1; GSPDB:GN00
A:Experimental source: strain C-125
C:Genetics:
A:Gene: aprX

Query Match 37.3%; Score 47; DB 2; Length 444;
Best Local Similarity 47.1%; Pred. No. 28;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 2 DLOERGDNDISFGDG 18
DB 319 DTTREDDDVAFSSRG 335

RESULT 8

AH0410
global stress requirement protein Gsra [Imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 17-May-2002
C:Accession: AH0410

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibbitt, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,
Nature 413, 533-537, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AH0410

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-481 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC92612.1; PID:g15981308; GSPDB:GN00175
C:Genetics:
A:Gene: gsra

C:Superfamily: Helicobacter serine proteinase

Query Match 37.3%; Score 47; DB 2; Length 481;
Best Local Similarity 62.5%; Pred. No. 30;
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

OY 7 GDNDISPFSGDGPFRK 22
DB 80 GDN--SPFGDGSFPQ 93

RESULT 9

B89944
hypothetical protein SA1447 [Imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: B89944

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.;
ma, A.; Mizutani-H, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu,
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsuki, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:1141816
A:Accession: B89944

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-825 <KUR>
A:Cross-references: GB:BA000018; PID:g13701417; PIDN:BA042711.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA1447

Query Match 37.3%; Score 47; DB 2; Length 825;
Best Local Similarity 58.8%; Pred. No. 55;
Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 7 GDNDISPFSGDGPFRD 23
DB 476 GDEDDLPVSGQVFRD 492

RESULT 10

538783
Integrin alpha chain - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 22-Jan-1994 #sequence_revision 14-Jul-1995 #text_change 29-Sep-1999
C:Accession: S38783; S23600
R:Kaufman, S.J.

submitted to the EMBL Data Library, March 1992
A:Reference number: S38783
A:Accession: S38783
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1106 <KAU>
A:Cross-references: EMBL:X65036; NID:g56392; PIDN:CAA6170.1; PID:g56393
J:Song, W.K.; Wang, W.; Foster, R.F.; Bleser, D.A.; Kaufman, S.J.
J. Cell Biol. 117, 643-657, 1992
A:Title: H36-alpha7 is a novel integrin alpha chain that is developmentally regulated
A:Reference number: S23600; MUID:92242309; PMID:1315319
A:Accession: S23600

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-116, 'R', 118-349, 'D', 351-804, 'R', 806, 'V', 808-1106 <SON>
A:Cross-references: EMBL:X65036
C:Superfamily: Integrin alpha-2b chain

Query Match 36.9%; Score 46.5; DB 2; Length 1106;
Best Local Similarity 47.8%; Pred. No. 91;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

OY 2 DLOERGDNDI---SPFGDGPFR 21
DB 362 DLNDGPPDIAGAPFDGKRVF 384

RESULT 11

161186
alpha-7 integrin - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
C:Accession: 161186

R:Zlobner, B.L.; Vu, M.P.; Maleh, N.; Crawford, J.; Lin, C.S.; Krimer, R.H.
J. Biol. Chem. 268, 26773-26783, 1993
A:Title: Alternative extracellular and cytoplasmic domains of the integrin alpha 7 su
A:Reference number: A49691; MUID:94075378; PMID:8253814
A:Accession: 161186
A:Status: preliminary; translated from GB/EMBL/DBJ

Fri Jan 17, 13:11:06 2003

us-09-641-034-47.rpr

Page 5

```

A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: AB3170
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-347 <5'3'>
A:Cross-references: GB:AE004799; GB:AE004091; NID:99949981; PIDN:AGC07211.1; GSPDB:GN001
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: queA; PA3824
C:Superfamily: Escherichia coli trna ribosyltransferase-isomerase
C:Keywords: Intramolecular transferase; isomerase

Query Match          36.1%; Score 45.5; DB 2; Length 347;
Best Local Similarity 42.9%; Pred. No. 36;
Matches 9; Conservative 3; Mismatches 2; Indels 7; Gaps 1;

Oy      8 DNDISPFSGD-----GQPF 21
      |::|||||               |::||
Db      263 DGEIKPFSGDIDFIYKGRPF 283

```

Search completed: January 17, 2003, 13:17:03
Job time : 18 secs

Fri Jan 17 13:11:06 2003

us-09-641-034-47.rsp

Page 1

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 13:16:38 ; Search time 10 Seconds

(without alignments)
95.396 Million cell updates/sec

Title: US-09-641-034-47

Sequence: 1 TDQERGDNDISPFSGDQPFKD 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	41.3	301	1	P31243 rhodobacter
2	49.5	39.3	1259	1	Q05695 rattus norv
3	48	38.1	385	1	P41453 autographa
4	46.5	36.9	1106	1	O63258 rattus norv
5	46.5	36.9	1179	1	O61738 mus musculu
6	46.5	36.9	1181	1	O36883 homo sapien
7	45.5	36.1	1015	1	O37474 drosophila
8	45.5	36.1	1260	1	P11627 mus musculu
9	45.5	36.1	1260	1	P11627 mus musculu
10	45	35.7	301	1	TSAL2_MOUSE
11	45	35.7	309	1	TSAL2_MOUSE
12	45	35.7	333	1	GUNI_USYMA
13	45	35.7	472	1	SACB_BACAM
14	45	35.7	508	1	VLI_HPV60
15	45	35.7	851	1	YDBG_SCHPO
16	44.5	35.3	628	1	YDBG_SCHPO
17	44.5	35.3	1050	1	ITAB_XENIA
18	44	34.9	475	1	NTCI_CRIGR
19	44	34.9	475	1	DEGP_SALTY
20	43.5	34.5	284	1	RL2_CHLNU
21	43.5	34.5	284	1	RL2_CHLNU
22	43.5	34.5	631	1	COBT_PSEDE
23	43.5	34.5	1053	1	ITAB_MOUSE
24	43.5	34.5	1115	1	ITAB_MOUSE
25	43	34.1	209	1	COEC_SCYCA
26	43	34.1	684	1	RPSD_AGRFS
27	43	34.1	3703	1	ABFI_HUMAN
28	43	34.1	3726	1	ABFI_HUMAN
29	42.5	33.7	488	1	FAIO_HUMAN
30	42.5	33.7	1000	1	ITAB_MOUSE
31	42	33.3	119	1	GVOI_HALNI
32	42	33.3	120	1	GVOI_HALNI
33	42	33.3	224	1	DCL_LICES

34	42	33.3	269	1	MHPD_ECOLI	P77608 escherichia
35	42	33.3	359	1	HRCA_RHIME	O92SKI rhizobium m
36	42	33.3	445	1	SLP2_DROME	P32031 drosophila
37	42	33.3	557	1	FLIE_RHIME	O54239 rhizobium m
38	42	33.3	823	1	SYL_DEIRA	O958FO delinococcus
39	42	33.3	1192	1	METH_MYCTU	O33259 mycobacteri
40	42	33.3	1756	1	TRIL_ECOLI	P14565 escherichia
41	41.5	32.9	232	1	VHED_BPT3	P20313 bacterioph
42	41.5	32.9	265	1	FAEH_ECOLI	P33782 escherichia
43	41.5	32.9	318	1	RBSK_SCHPO	O60116 schizosacch
44	41.5	32.9	811	1	ENAN_BPKIE	P49714 bacterioph
45	41.5	32.9	1025	1	ITAB_HUMAN	P53708 homo sapien

ALIGNMENTS

```
RESULT 1
PORI_RHOCA
ID AC P31243:
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Porin.
OS Rhodobacter capsulatus (Rhodospseudomonas capsulata).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
CC Rhodobacter.
OX NCBI_TaxID=1061;
RN [1]
RP SEQUENCE.
RC STRAIN-DSM 938 / 37B4;
RX MEDLINE-9130909; PubMed-1651239;
RA Schiltz E., Kreusch A., Nestel U., Schulz G.E.;
RT "Primary structure of porin from Rhodobacter capsulatus.";
RL Eur. J. Biochem. 199;587-594(1991).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RC STRAIN-DSM 938 / 37B4;
RX MEDLINE-9036791; PubMed-2165921;
RA Weiss M.S., Wacker T., Weckesser J., Walte W., Schulz G.E.;
RT "The three-dimensional structure of porin from Rhodobacter capsulatus at 3-A resolution.";
RL FEBS Lett. 267:268-272(1990).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
RC STRAIN-DSM 938 / 37B4;
RX MEDLINE-91192174; PubMed-1707373;
RA Weiss M.S., Kreusch A., Schiltz E., Nestel U., Walte W., Weckesser J., Schulz G.E.;
RT "The structure of porin from Rhodobacter capsulatus at 1.8-A resolution.";
RL FEBS Lett. 280:379-382(1991).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
RX MEDLINE-93021091; PubMed-1328651;
RA Weiss M.S., Schulz G.E.;
RT "Structure of porin refined at 1.8-A resolution.";
RL J. Mol. Biol. 227:493-509(1992).
CC -1- FUNCTION: FORMS CHANNELS THAT ALLOW THE PASSIVE DIFFUSION OF SMALL HYDROPHILIC SOLUTIONS UP TO AN EXCLUSION LIMIT OF ABOUT 0.6 kDa.
CC -1- SUBUNIT: HOMOTRIMER.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane. PEATED SHEET BARREL (171 RESIDUES), THREE SHORT ALPHA-HELICES (18 RESIDUES) AND 13 HYDROGEN-BONDED REVERSE TURNS (26 RESIDUES).
CC PIR: S16070; S16070.
CC PDB: 2POR; 15-JUL-93.
CC Outer membrane; Transmembrane; Porin; 3D-structure.
FT STRAND 2 14
FT STRAND 19 34
FT TURN 36 37
```

[illegible]

FT CARBOHYD 1106 1106 N-LINKED (GLCNAC...) (POTENTIAL).
 FT VARSPLIC 1179 1182 MISSING (IN ISOFORM 2).
 SQ SEQUENCE 1259 AA: 140934 MW: 0F12A7C4415F3C08 CRC64;
 Query Match 39.3%; Score 49.5; DB 1; Length 1259;
 Best Local Similarity 66.7%; Pred. NO. 13;
 Matches 12; Conservative 1; Mismatches 4; Indels 1; Gaps 1;
 QY 2 DLQERGDNDISFGDQ 19
 DB 558 DLQERGDSD-KYFIEDQ 574
 RESULT 3
 LEF3_NPVAC STANDARD: PRT: 385 AA.
 AC P41453;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Late expression factor 3.
 GN LEF-3.
 OS Autographa californica nuclear polyhedrosis virus (AcMNPV).
 OC Viruses: dsDNA viruses, no RNA stage; Baculoviridae;
 CC Nucleopolyhedrovirus.
 CX NCBI_TaxID=46015;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-L1;
 RX MEDLINE=93353600; PubMed=8350397;
 RA Li Y., Passarelli A.L., Miller L.K.;
 RT "Identification, sequence, and transcriptional mapping of lef-3, a
 baculovirus gene involved in late and very late gene expression.";
 RL J. Virol. 67:5260-5268(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C6;
 RX MEDLINE=94303173; PubMed=8030224;
 RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
 RT "The complete DNA sequence of Autographa californica nuclear
 polyhedrosis virus.";
 RL Virology 202:586-605(1994).
 CC -1- FUNCTION: REQUIRED FOR LATE AND VERY LATE GENE EXPRESSION. LEF-3
 COULD BE A SINGLE STRANDED DNA-BINDING PROTEIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: L18873; AAA02964.1;
 DR EMBL: L22858; AAA66697.1;
 DR PIR: A40677; A40677.
 KW Early protein; transcription regulation; DNA-binding.
 SQ SEQUENCE 385 AA: 44551 MW: 9A25ECD7BA7EBDF1 CRC64;
 Query Match 38.1%; Score 48; DB 1; Length 385;
 Best Local Similarity 56.2%; Pred. NO. 5.5;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 QY 8 DNDISFGDQGFQD 23
 DB 61 DNKIQEYGVGDSQSFQD 76
 RESULT 4
 ITA7_RAT STANDARD: PRT: 1106 AA.
 AC 063258; 063026; 063027;
 DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Integrin alpha-7 (H36-alpha7).
 GN ITGA7.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM ALPHA-7X1B).
 RC TISSUE-Skeletal muscle;
 RX MEDLINE=92242309; PubMed=1315319;
 RA Song W.K., Wang W., Foster R.F., Bleiser D.A., Kaufman S.J.;
 RT "H36-alpha 7 is a novel integrin alpha chain that is developmentally
 RT regulated during skeletal myogenesis.";
 RL J. Cell Biol. 117:643-657(1992).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS ALPHA-7X1A AND ALPHA-7X1C).
 RC TISSUE-Skeletal muscle;
 RX MEDLINE=94171924; PubMed=8126096;
 RA Song W.K., Wang W., Sato H., Bleiser D.A., Kaufman S.J.;
 RT "Expression of alpha 7 integrin cytoplasmic domains during skeletal
 RT muscle development: alternate forms, conformational change, and
 RT homologies with serine/threonine kinases and tyrosine phosphatases.";
 RL J. Cell Sci. 106:1139-1152(1993).
 RN [3]
 RP TISSUE SPECIFICITY.
 RX MEDLINE=96197133; PubMed=8626012;
 RA Martin P.T., Kaufman S.J., Kramer R.H., Sanes J.R.;
 RT "Synaptic integrins in developing, adult, and mutant muscle: selective
 RT association of alpha1, alpha7A, and alpha7B integrins with the
 RT neuromuscular junction.";
 RL Dev. Biol. 174:125-139(1996).
 CC -1- FUNCTION: INTEGRIN ALPHA-7/BETA-1 IS THE PRIMARY LAMININ RECEPTOR
 CC ON SKELETAL MYOBLASTS AND ADULT MYOFIBERS. DURING MYOGENIC
 CC DIFFERENTIATION, IT MAY INDUCE CHANGES IN THE SHAPE AND MOBILITY
 CC OF MYOBLASTS, AND FACILITATE THEIR LOCALIZATION AT LAMININ-RICH
 CC SITES OF SECONDARY FIBER FORMATION. INVOLVED IN THE MAINTENANCE OF
 CC THE MYOFIBERS CYTOARCHITECTURE AS WELL AS FOR THEIR ANCHORAGE,
 CC VIABILITY AND FUNCTIONAL INTEGRITY.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
 CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAIN LINKED BY A
 CC DISULFIDE BOND. ALPHA-7 ASSOCIATES WITH BETA-1.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS: ALPHA-7X1A, ALPHA-7X1B
 CC (SHOWN HERE) AND ALPHA-7X1C; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN SKELETAL AND CARDIAC MUSCLE.
 CC EXPRESSED IN REPLICATING MYOBLASTS. IN DIFFERENTIATED MUSCLE
 CC FIBERS LOCALIZES BETWEEN FIBERS AND THE SURROUNDING MATRIX.
 CC ISOFORMS A AND B ARE EXPRESSED AT MYOTENDINUS AND NEUROMUSCULAR
 CC JUNCTIONS; ISOFORM C IS EXPRESSED AT NEUROMUSCULAR JUNCTIONS AND
 CC AT EXTRASYNAPTIC SITES.
 CC -1- DEVELOPMENTAL STAGE: ISOFORMS ARE DEVELOPMENTALLY REGULATED DURING
 CC THE FORMATION OF SKELETAL MUSCLE. ISOFORMS A AND C ARE INDUCED
 CC UPON TERMINAL MYOGENIC DIFFERENTIATION; ISOFORM B IS PRESENT
 CC EARLIER IN REPLICATING CELLS AND DIMINISHES UPON DIFFERENTIATION.
 CC -1- PTM: ADP-RIBOSYLATED ON AT LEAST TWO SITES IN THE EXTRACELLULAR
 CC DOMAIN IN SKELETAL MYOTUBES (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
 CC -1- SIMILARITY: CONTRAINS 7 PG-GAP REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: X65036; CAA46170.1;
 DR EMBL: X74293; CAA52346.1;
 DR EMBL: X74294; CAA52347.1;
 DR HSP: P11215; IABX.

RT subunit generated by alternative intron splicing.";
 RL Biochem. Biophys. Res. Commun. 243:317-325(1998).
 [12]
 RP SEQUENCE FROM N.A. (ISOFORM ALPHA-7X2B).
 RX MEDLINE-98250181; PubMed-9590299;
 RA Hayashi Y.K., Chou F.-L., Engvall E., Ogawa M., Matsuda C.,
 RA Hirabayashi S., Yokochi K., Zieber B.L., Kramer R.H., Kaufman S.J.,
 RA Oawa E., Goto Y.-T., Nonaka I., Tsukahara T., Wang J.Z.,
 RA Hoffman E.P., Arakata K.;
 RT "Mutations in the integrin alpha7 gene cause congenital myopathy.";
 RL Nat. Genet. 19:94-97(1998).
 [13]
 RP SEQUENCE FROM N.A. (ISOFORM ALPHA-7X2B).
 RA Viziianakis I.S., Zieber B.L., Kramer R.H.;
 RT "Cloning of human integrin alpha-7 cDNA.";
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 [14]
 RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
 RC TISSUE-Skeletal muscle;
 RX MEDLINE-99333684; PubMed-10403775;
 RA Vignier N., Moghadassadeh B., Gary F., Beckmann J., Mayer U.,
 RA Gulcener P.;
 RT "Structure, genetic localization, and identification of the cardiac
 RT and skeletal muscle transcripts of the human integrin alpha7 gene
 RT (ITGA7).";
 RL Biochem. Biophys. Res. Commun. 260:357-364(1999).
 [15]
 RP SEQUENCE OF 34-45.
 RC TISSUE-Melanoma;
 RX MEDLINE-92198982; PubMed-1839357;
 RA Kramer R.H., Vu M.P., Cheng Y.F., Ramos D.M., Timpl R., Waleh N.;
 RT "Laminin-binding integrin alpha 7 beta 1: functional characterization
 RT and expression in normal and malignant melanocytes.";
 RL Cell Regul. 2:805-817(1991).
 [16]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORM ALPHA-7X2).
 RC TISSUE-Heart;
 RX MEDLINE-94075378; PubMed-8253814;
 RA Zieber B.L., Vu M.P., Waleh N., Crawford J., Lin C.-S., Kramer R.H.;
 RT "Alternative extracellular and cytoplasmic domains of the integrin
 RT alpha 7 subunit are differentially expressed during development.";
 RL J. Biol. Chem. 268:26773-26783(1993).
 [17]
 RP SEQUENCE OF 1105-1181 FROM N.A. (ISOFORM ALPHA-7B).
 RC TISSUE-Fetal muscle;
 RX MEDLINE-941171924; PubMed-8126096;
 RA Song W.K., Wang W., Sato H., Bielese D.A., Kaufman S.J.;
 RT "Expression of alpha 7 integrin cytoplasmic domains during skeletal
 RT muscle development: alternate forms, conformational change, and
 RT homologues with serine/threonine kinases and tyrosine phosphatases.";
 RL J. Cell Sci. 106:1139-1152(1993).
 [18]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORM ALPHA-7A).
 RC TISSUE-Skeletal muscle;
 RX MEDLINE-98012902; PubMed-9352853;
 RA Basora N., Vachon P.H., Herring-Gilliam F.E., Perreault N.,
 RA Beaulieu J.-F.;
 RT "Relation between integrin alpha7beta1 expression in human intestinal
 RT cells and enterocytic differentiation.";
 RL Gastroenterology 113:1510-1521(1997).
 [19]
 RP FUNCTION.
 RX MEDLINE-97453229; PubMed-9307969;
 RA Zieber B.L., Chen Y.Q., Kramer R.H.;
 RT "The laminin-binding activity of the alpha 7 integrin receptor is
 RT defined by developmentally regulated splicing in the extracellular
 RT domain.";
 RL Mol. Biol. Cell 8:1723-1734(1997).
 [110]
 RP FUNCTION.
 RX MEDLINE-20160722; PubMed-10694445;
 RA Schreiber S., Mehlitz D., Echterneyer F., Hapke S., Poeschl E.,
 RA von der Mark H., Moch H., von der Mark K.;

RT "The role of extracellular and cytoplasmic splice domains of
 RT alpha7-integrin in cell adhesion and migration on laminins.";
 RL Exp. Cell Res. 255:303-313(2000).
 [111]
 RP TISSUE SPECIFICITY.
 RX MEDLINE-96197133; PubMed-8626012;
 RA Martin P.T., Kaufman S.J., Kramer R.H., Sanes J.R.;
 RT "Synaptic integrin in developing, adult, and mutant muscle: selective
 RT association of alpha1, alpha7A, and alpha7B integrins with the
 RT neuromuscular junction.";
 RL Dev. Biol. 174:125-139(1996).
 [112]
 RP FUNCTION: INTEGRIN ALPHA-7/BETA-1 IS THE PRIMARY LAMININ RECEPTOR
 ON SKELETAL MYOBLASTS AND ADULT MYOFIBERS. DURING MYOGENIC
 DIFFERENTIATION, IT MAY INDUCE CHANGES IN THE SHAPE AND MOBILITY
 OF MYOBLASTS, AND FACILITATE THEIR LOCALIZATION AT LAMININ-RICH
 CC SITES OF SECONDARY FIBER FORMATION. IT IS INVOLVED IN THE
 CC MAINTENANCE OF THE MYOFIBERS CYTOARCHITECTURE AS WELL AS FOR THEIR
 CC ANCHORAGE, VIABILITY AND FUNCTIONAL INTEGRITY. ISOFORMS ALPHA-7X2B
 CC AND ALPHA-7X1B PROMOTE MYOBLAST MIGRATION ON LAMININ 1 AND LAMININ
 CC 2/4, BUT ALPHA-7X1B IS LESS ACTIVE ON LAMININ 1 (IN VITRO).
 CC SUBUNIT IS COMPOSED OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
 CC DISULFIDE BOND. ALPHA-7 ASSOCIATES WITH BETA-1.
 CC [113]
 RP SUBCELLULAR LOCATION: Type I membrane protein.
 CC [114]
 RP ALTERNATIVE PRODUCTS: AT LEAST 12 ISOFORMS: ALPHA-7X1A, ALPHA-
 CC 7X1B, ALPHA-7X1DA, ALPHA-7X1DB, ALPHA-7X2A, ALPHA-7X2B, ALPHA-
 CC 7X2DA, ALPHA-7X2DB, ALPHA-7X1X2A, ALPHA-7X1X2B (SHOWN HERE),
 CC ALPHA-7X1X2DA AND ALPHA-7X1X2DB, MAY BE PRODUCED BY ALTERNATIVE
 CC SPLICING. THERE IS A COMBINATION OF AT LEAST FIVE ALTERNATIVE
 CC SPLICED DOMAINS, THREE EXTRACELLULAR (X1, X2 AND D) AND TWO
 CC CYTOPLASMIC (A AND B). A THIRD POTENTIAL ALTERNATIVE SPLICED
 CC CYTOPLASMIC DOMAIN (C) DOES NOT APPEAR TO BE EXPRESSED. IN VITRO
 CC GENERATED ISOFORM X2C SHOWS FUNCTION. SO FAR DETECTED ARE ISOFORMS
 CC ALPHA-7X1A, ALPHA-7X2B (MAJOR), ALPHA-7X2DB (MINOR) AND MINOR
 CC ISOFORMS CONTAINING SEGMENT X1X2.
 [115]
 RP TISSUE SPECIFICITY: ISOFORMS CONTAINING SEGMENT A ARE
 CC PREDOMINANTLY EXPRESSED IN SKELETAL MUSCLE. ISOFORMS CONTAINING
 CC SEGMENT B ARE ABUNDANTLY EXPRESSED IN SKELETAL MUSCLE, MODERATELY
 CC IN CARDIAC MUSCLE, SMALL INTESTINE, COLON, OVARY AND PROSTATE AND
 CC WEAKLY IN LUNG AND TESTES. ISOFORMS CONTAINING SEGMENT X2D ARE
 CC EXPRESSED AT LOW LEVELS IN FETAL AND ADULT SKELETAL MUSCLE AND IN
 CC CARDIAC MUSCLE, BUT ARE NOT DETECTED IN MYOBLASTS AND MYOTUBES. IN
 CC MUSCLE FIBERS ISOFORMS CONTAINING SEGMENT A AND B ARE EXPRESSED AT
 CC MYOTENDINOUS AND NEUROMUSCULAR JUNCTIONS. ISOFORMS CONTAINING
 CC SEGMENT C ARE EXPRESSED AT NEUROMUSCULAR JUNCTIONS AND AT
 CC EXTRASYNAPTIC SITES. ISOFORMS CONTAINING SEGMENTS X1 OR X2 OR, AT
 CC LOW LEVELS, X1X2 ARE EXPRESSED IN FETAL AND ADULT SKELETAL MUSCLE
 CC (MYOBLASTS AND MYOTUBES) AND CARDIAC MUSCLE.
 [116]
 RP DEVELOPMENTAL STAGE: IN RENOVING INTESTINAL EPITHELIUM, EXPRESSION
 CC OF ISOFORMS CONTAINING SEGMENT B CORRELATES WITH THE ONSET OF
 CC ENTEROCYTIC DIFFERENTIATION.
 [117]
 RP PTM: ADP-RIBOSYLATED ON AT LEAST TWO SITES OF THE EXTRACELLULAR
 CC DOMAIN IN SKELETAL MYOTUBES (BY SIMILARITY).
 [118]
 RP DISEASE: DEFECTS IN ITGA7 ARE ASSOCIATED WITH A FORM OF CONGENITAL
 CC MYOPATHY, A GROUP OF HETEROGENEOUS MUSCLE DISORDERS WHICH ARE
 CC THOUGHT TO RESULT FROM ABNORMAL MUSCLE DEVELOPMENT. MUSCLE
 CC WEAKNESS IS EITHER NON-PROGRESSIVE OR SLOWLY PROGRESSIVE AND
 CC APPARENT FROM BIRTH OR EARLY INFANCY.
 [119]
 RP SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
 [120]
 RP SIMILARITY: CONTAINS 7 FG-GAP REPEATS.
 [121]
 RP THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
 CC BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -
 CC THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
 CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
 CC MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
 CC ENTITIES REQUIRES A LICENSE AGREEMENT (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.slb.ch).
 [122]
 RP EMBL: AF032108; AAC39708.1;
 DR EMBL: AF052050; AAC18968.1;
 DR EMBL: AF072132; AAC80458.1;
 DR EMBL: AJ228836; CAB41534.1;


```

CC -1- FUNCTION: Synthesizes oQ from preQ1 in a single S-
CC adenosylmethionine-requiring step. The ribosyl moiety of adomet is
CC transferred and isomerized to the epoxycyclopentane residue of oQ
CC (by similarity).
CC -1- PATHWAY: Queuosine biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (potentia1).
CC -1- SIMILARITY: BELONGS TO THE QUEA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
CC -----
DR EMBL; AE004799; AAC07211.1; -
DR InterPro: IPR003699; Queuosine_synth.
DR Pfam: PF02547; Queuosine_synth; 1.
DR TIGRFAMs: TIGR00113; quea; 1.
DR K09 Quenosine biosynthesis; Transferase; Isomerase; Complete proteome.
SQ SEQUENCE 347 AA; 38160 MM; 48ECBF074C2BC589 CRC64;

Query Match 36.1% Score 45.5; DB 1; Length 347;
Best Local Similarity 42.9% Pred. No. 12;
Matches 9; Conservative 3; Mismatches 2; Indels 7; Gaps 1;

Db 263 DGEIKPFGSDHDTIRYPERP 283

      8 DNDISPFSGD-----GQPF 21
      |:::||||| 1::|
      263 DGEIKPFGSDHDTIRYPERP 283

RESULT 8
ITAA_DROME
ID ID ITAA_DROME STANDARD: PRT: 1015 AA.
AC Q9Y7A4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Integrin alpha-PS4 precursor (Position-specific antigen 4, alpha
DE chain).
DE CG16827.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RF STRAIN=Berkelley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Arroll J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cleyser S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson R., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evansglista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Posler C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Harris A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush G., Karpen G.H., Ke Z., Kennison J.B., Ketchum K.A.,
RA Kimmel B., Klotz C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lascko P., Leal V., Leavitt J.A., Li J., Lin S., Little J., Liu X.,

```



```

FT DISULFID 157 208 BY SIMILARITY.
FT DISULFID 263 311 BY SIMILARITY.
FT DISULFID 353 403 BY SIMILARITY.
FT DISULFID 447 496 BY SIMILARITY.
FT DISULFID 538 590 BY SIMILARITY.
FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 202 202 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 478 478 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 504 504 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 587 587 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 670 670 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 725 725 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 776 776 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 824 824 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 848 848 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 875 875 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 968 968 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 978 978 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1022 1022 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1030 1030 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1073 1073 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1107 1107 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1260 AA: 140968 MW: 22857001CB2A538 CRC64:

```

```

Query Match 36.1% Score 45.5: DB 1: Length 1260:
Best Local Similarity 61.1% Pred. No. 53:
Matches 11: Conservative 2; Mismatches 4; Indels 1; Gaps 1;
OY 2 DLQERGNDISPFGDQ 19
Db 558 DLQERGNDS-KYFIEDCK 574

```

```

RESULT 10
TS2A_MOUSE STANDARD: PRT: 301 AA.
AC OGVIG3: O9DAL5:
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Testis-specific gene A2 (Male meiotic metaphase chromosome-associated
  acidic protein) (Melchroacidin).
GN TSGA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A. AND CHARACTERIZATION.
RC TISSUE-Testis:
RX MEDLINE-98246622: Pubmed-9578619;
RA Tsuchida J., Nishina Y., Wakabayashi N., Nozaki M., Sakai Y.,
RA Nishimune Y.;
RT "Molecular cloning and characterization of melchroacidin (male meiotic
  metaphase chromosome-associated acidic protein).";
RL Dev. Biol. 197:67-76(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL/6J: TISSUE-Testis:
RX MEDLINE-21085660: Pubmed-11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishikawa K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gliss C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,

```

```

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guinacino S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Marzetti J., Mombert P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Suzuki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Hayashizaki Y.;
RA Wyszewski A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- FUNCTION: The specific expression during male germ cell
CC development and its characteristic localization suggest that it
CC may play an important role in male meiosis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic in late spermatocytes, secondary
CC spermatocytes and round spermatids. Gathered around metaphase
CC chromosomes during meiotic divisions.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; may be
CC produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: Germ cell specific. Specifically expressed in
CC testis, and to a lower extent in ovary. Not expressed in somatic
CC tissues.
CC -1- DEVELOPMENTAL STAGE: During male germ cell development it is not
CC detected until 12 days. Significant expression is detected from
CC 14-day-old through to adult testis. Expression is first detected
CC in the pachytene spermatocytes at stage V, becomes stronger from
CC the late pachytene spermatocytes to round spermatid stage, and
CC then gradually decreases as the morphogenesis proceeds further.
CC Not expressed in germ cells located in the first layer of the
CC seminiferous epithelium (spermatogonia, leptotene and zygotene
CC spermatocytes).
CC -1- SIMILARITY: CONTAINS 6 MORN REPEATS.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 283.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AB005635; BAB83693.1; ALT_FRAME.
DR MGD: AK005739; BAB24214.1;
DR MGD: MGI:1194909; Tsga2.
DR InterPro: IPR003409; MORN.
DR Pfam: PF02493; MORN; 6.
DR KEGG: Repeat; Repeat; Alternative splicing.
FT DOMAIN 195 201 POLY-GLU.
FT REPEAT 20 43 MORN REPEAT 1.
FT REPEAT 44 66 MORN REPEAT 2.
FT REPEAT 67 89 MORN REPEAT 3.
FT REPEAT 90 112 MORN REPEAT 4.
FT REPEAT 113 135 MORN REPEAT 5.
FT REPEAT 159 181 MORN REPEAT 6.
FT REPEAT 181 181 MORN REPEAT 6.
FT REPEAT 181 181 MORN REPEAT 6.
SQ SEQUENCE 301 AA: 34181 MW: 704ABEFB94DFC90B CRC64:
Query Match 35.7% Score 45: DB 1: Length 301:
Best Local Similarity 37.5% Pred. No. 12:
Matches 6: Conservative 6; Mismatches 4; Indels 0; Gaps 0;
OY 2 DLQERGNDISPFGD 17
Db 8 ELKEEGNDLCEGE 23

```

```

RESULT 11
TS2A_HUMAN STANDARD: PRT: 309 AA.
AC OGVIR4:
DT 15-JUN-2002 (Rel. 41, Created)

```

DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Testis-specific gene A2 (Male meiotic metaphase chromosome-associated
 DE acidic protein) (Melchroscidin).
 GN TSGA2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE-Testis;
 RX MEDLINE=98246622; PubMed=9578619;
 RA Tsuchida J., Nishina Y., Wakabayashi N., Nozaki M., Sakai Y.,
 RA Nishimune Y.;
 RT Molecular cloning and characterization of melchroscidin (male meiotic
 RT metaphase chromosome-associated acidic protein)."
 RL Dev. Biol. 197;67-76(1998).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC TISSUE-Testis;
 RA Shimizu N., Kudoh J., Shibuya K.;
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: May play an important role in male meiosis (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic. In late spermatocytes, secondary
 CC chromosomes during meiotic divisions. Gathered around metaphase
 CC -1- TISSUE SPECIFICITY: Testis specific.
 CC -1- SIMILARITY: CONTAINS 6 MORN REPEATS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: AB06536; BAB83694.1; -
 DR EMBL: AB041016; BAB92995.1; -
 DR Genev; HGNC:12371; TSGA2.
 DR InterPro: IPR003409; MORN.
 DR Pfam; PF02493; MORN; 6.
 DR
 KW Melosis; Repeat.
 FT DOMAIN 195 201 POLY-GLU.
 FT 301 301 POLY-GLU.
 FT REPEAT 20 43 MORN REPEAT 1.
 FT REPEAT 44 66 MORN REPEAT 2.
 FT REPEAT 67 89 MORN REPEAT 3.
 FT REPEAT 90 112 MORN REPEAT 4.
 FT REPEAT 113 135 MORN REPEAT 5.
 FT REPEAT 159 181 MORN REPEAT 6.
 FT SEQUENCE 309 AA; 35124 MW; 09A20A76D5A6AB8 CRC64;
 SQ
 Query Match 35.7%; Score 45; DB 1; Length 309;
 Best Local Similarity 46.7%; Pred. No. 12;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 QY 2 DQEGDNDISPFG 16
 DB 8 ELEEGENDIGEYEG 22
 DB
 RESULT 12
 GUN1_USTMA STANDARD; PRT; 393 AA.
 ID GUN1_USTMA
 AC P54424;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Endoglucanase I precursor (EC 3.2.1.4) (Endo-1,4-beta-glucanase 1)
 DE (Cellulase 1) (EG 1).

GN EGLI.
 OS Ustilago maydis (Smut fungus).
 OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
 OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
 OX NCBI_TaxID=5270;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=PD11;
 RX MEDLINE=96145728; PubMed=8590631;
 RA Schaevecker F., Wanner G., Kahmann R.;
 RT Filament-specific expression of a cellulase gene in the dimorphic
 RT fungus Ustilago maydis."
 RL Biol. Chem. Hoppe-Seyler 376:617-625(1995).
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
 CC linkages in cellulose.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: HYPHAL TIP.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN FILAMENTOUS DIKARYON.
 CC -1- PTM: MAY ALSO BE O-GLYCOSYLATED.
 CC -1- SIMILARITY: BELONGS TO CELLULASE FAMILY K (FAMILY 45 OF GLYCOSYL
 CC HYDROLASES).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: S81598; AAB36147.1; -
 DR HSSP; P43316; ZENG.
 DR InterPro: IPR000334; GH_45.
 DR Pfam; PF02015; Glyco_hydro_45; 1.
 DR PROSITE; PS01140; GLYCOSYL_HYDROL_F45; 1.
 KW Cellulose degradation; Hydrolase; Glycosidase; Signal; Glycoprotein.
 FT CHAIN 1 26 POTENTIAL.
 FT 27 393 ENDOGLUCANASE 1.
 FT ACT_SITE 34 34 NUCLEOPHILE (BY SIMILARITY).
 FT ACT_SITE 152 152 PROTON DONOR (BY SIMILARITY).
 FT DOMAIN 270 385 ALA/GLY/SER-RICH.
 FT CARBOHYD 343 343 N-LINKED (GLCNAC... (POTENTIAL).
 FT SEQUENCE 393 AA; 39594 MW; 65C753C610CD6AD3 CRC64;
 SQ
 Query Match 35.7%; Score 45; DB 1; Length 393;
 Best Local Similarity 61.5%; Pred. No. 16;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 4 QEGDNDISPFG 16
 DB 235 QRRDNDTISPSG 247
 DB
 RESULT 13
 SABC_BACAM
 ID SABC_BACAM STANDARD; PRT; 472 AA.
 AC P21130;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE levansucrase precursor (EC 2.4.1.10) (Beta-D-fructofuranosyl
 DE transferase) (Sucrose 6-fructosyl transferase).
 GN SABC.
 OS Bacillus amyloliquefaciens.
 OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1390;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 23844;
 RX MEDLINE=91092506; PubMed=2265762;
 RA Tang L.B., Lenstra R., Borchert T.V., Vasantha N.;
 RT Isolation and characterization of levansucrase-encoding gene from
 RT Bacillus amyloliquefaciens."

RL Gene 96:89-93(1990).
 CC -1- CATALYTIC ACTIVITY: Sucrose + ((2,6)-beta-D-fructosyl)(N) -
 CC GLUCOSE + ((2,6)-beta-D-fructosyl)(N+1).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- INDUCTION: BY SUCROSE.
 CC -1- SIMILARITY: BELONGS TO FAMILY 68 OF GLYCOSYL HYDROLASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: X52988; CAA37179.1; -
 CC PIR: J00802; J00802.
 CC PIR: S11739; S11739.
 CC InterPro: IPR003469; Glyco_hydro_68.
 CC Pfam: PF02435; Glyco_hydro_68; 1.
 CC Transferase: Glycosyltransferase; Signal.
 CC SIGNAL 1
 CC CHAIN 30 472 LEVANSUCRASE.
 CC FT
 CC SEQUENCE 472 AA; 52859 MW; F38592D27267E7D CRC64;
 SO
 Query Match 35.7%; Score 45; DB 1; Length 472;
 Best Local Similarity 45.0%; Pred. No. 20;
 Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
 Oy 4 QERGNDISFSGDGFCD 23
 Db 126 QKVGNDISDMKMGKRVFED 145
 ID VLI_HPV60 STANDARD; PRT; 508 AA.
 AC P50821; O80947;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Major capsid protein LI.
 GN LI
 OS Human papillomavirus type 60.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OC NCBI_TaxID=40540;
 OX
 RN
 RN SEQUENCE FROM N.A.
 RA Delius H.;
 RN Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 372-467 FROM N.A.
 RA Chan S.Y., Delius H., Halpern A.L., Bernard H.U.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: U31792; AAA79491.1; -
 CC EMBL: U21876; AAA92837.1; -
 CC InterPro: IPR002210; PV_capsid_L1.
 CC Pfam: PF00500; late_protein_L1; 1.
 CC PRINTS: PR00865; HPVcapsidL1.
 CC ProDom: PD000544; PV_capsid_L1; 1.
 CC Coa protein; Late protein.
 CC SEQUENCE 508 AA; 57827 MW; 3149549895534D00 CRC64;
 SO

Query Match 35.7%; Score 45; DB 1; Length 508;
 Best Local Similarity 64.3%; Pred. No. 22;
 Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 Oy 2 DLQERNDISFSGD 15
 Db 459 DLQERFNSLSQFS 472
 ID YDQG_SCHPO STANDARD; PRT; 851 AA.
 AC Q10366;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein C22E12.16c in chromosome I.
 GN SPAC22E12.16c.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OC NCBI_TaxID=4896;
 OX
 RN
 RN SEQUENCE FROM N.A.
 RP STRAIN=972;
 RC MEDLINE=21848401; PubMed=11859360;
 RX Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skellton J., Simmonds M., Squares S., Stevens K.,
 RA Taylor K., Taylor R.C., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Wolchert G., Aert R., Robben J., Grynoprez B.,
 RA Weljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fitz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leinach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Medler H., Mambut R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McComble W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RA "The genome sequence of Schizosaccharomyces pombe".
 FT Nature 415:871-880(2002).
 CC
 CC -1- SIMILARITY: STRONG, TO YEAST PIK1.
 CC
 CC -1- SIMILARITY: BELONGS TO THE PI3/P14-KINASES FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: Z70043; CAA93903.1; -
 CC InterPro: IPR00403; PI3_P14_kinase.
 CC Pfam: PF00454; PI3_P14_kinase; 1.
 CC SMART: SM00146; PI3K; 1.
 CC PROSITE: PS00915; PI3_4_KINASE_1; 1.
 CC PROSITE: PS00916; PI3_4_KINASE_2; 1.
 CC PROSITE: PS02900; PI3_4_KINASE_3; 1.
 CC Hypothetical protein; Transferase; Kinase.
 CC DOMAIN 581 827 PI3K/PIAK.

SQ SEQUENCE 851 AA; 96657 MM; A991F3C7E3D980BE CRC64;
 Query Match 35.7%; Score 45; DB 1; Length 851;
 Best Local Similarity 50.0%; Pred. No. 40;
 Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 QY 1 TDLQERGDNDISPFGDG 18
 :||:||||:|
 Db 382 TNLQDSTDNDISESESEGC 399

Search completed: January 17, 2003, 13:18:02
 Job time : 11 secs

Gencore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 17, 2003, 13:16:38 ; Search time 29 Seconds

(without alignments)
163.417 Million cell updates/sec

Title: US-09-641-034-47

Sequence: 1 TDLQERGDNDISFGSDGQPKFD 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

1: SP archaea: *
2: SP bacteria: *
3: SP fungi: *
4: SP human: *
5: SP invertebrate: *
6: SP mammal: *
7: SP mhc: *
8: SP organelle: *
9: SP phage: *
10: SP plant: *
11: SP rodent: *
12: SP virus: *
13: SP vertebrate: *
14: SP unclassified: *
15: SP viirus: *
16: SP bacteriap: *
17: SP archaeap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	126	100.0	525	4	09N076
2	123	97.6	500	6	095K65
3	123	97.6	555	6	09GMI3
4	123	97.6	555	6	09N076
5	81	64.3	433	11	092411
6	81	64.3	433	11	092411
7	52	41.3	320	2	052676
8	49.5	39.3	793	5	09VYV0
9	49.5	38.9	752	4	09C091
10	48	38.1	277	16	08X183
11	48	38.1	385	12	092432
12	48	38.1	440	2	09F5V2
13	48	38.1	484	2	044497
14	48	38.1	481	5	09NK94
15	47	37.3	444	16	09KBJ7
16	47	37.3	477	2	08RSS1

17	47	37.3	481	16	08ZBM6	08ZBM6 yersinia pe
18	47	37.3	76	16	08XR29	08XR29 ralatonia s
19	47	37.3	825	16	0997N0	0997N0 staphylococ
20	46	36.5	279	2	055321	055321 synechococ
21	46	36.5	365	16	08U6F1	08U6F1 agrobacteri
22	45.5	36.1	265	2	033643	033643 salmonella
23	45.5	36.1	266	2	047220	047220 escherichia
24	45.5	36.1	1259	11	090Y38	090Y38 mus musculu
25	45	35.7	158	16	08Y4V1	08Y4V1 listeria mo
26	45	35.7	180	5	09NDR9	09NDR9 brachyosco
27	45	35.7	207	11	08V1K4	08V1K4 mus musculu
28	45	35.7	237	11	0912W8	0912W8 mus musculu
29	45	35.7	337	10	09S1M8	09S1M8 arabidopsis
30	45	35.7	536	12	08VA07	08VA07 chikungunya
31	45	35.7	690	4	096MB5	096MB5 homo sapien
32	45	35.7	913	16	09P003	09P003 ureaplasma
33	45	35.7	1052	2	09WHL5	09WHL5 methylolact
34	45	35.7	1419	5	09V1L3	09V1L3 sambucus n1
35	44.5	35.3	563	10	004367	004367 escherichia
36	44.5	35.3	698	16	08X4L8	08X4L8 rhizobium l
37	44	34.9	139	16	0983F0	0983F0 ralatonia s
38	44	34.9	156	16	08XUS9	08XUS9 aedes aegypt
39	44	34.9	215	5	0819T8	0819T8 neurospora
40	44	34.9	216	3	08X041	08X041 salmonella
41	44	34.9	475	16	08Z9B0	08Z9B0 plasmodium
42	44	34.9	533	5	094677	094677 plasmodium
43	44	34.9	572	5	09GSR0	09GSR0 plasmodium
44	44	34.9	682	2	09ACJ4	09ACJ4 pseudomonas
45	44	34.9	683	16	0914P3	0914P3 pseudomonas

ALIGNMENTS

RESULT 1
09N076 PRELIMINARY; PRT: 525 AA.
ID 09N076;
AC 09N076;
DT 01-OCT-2000 (TRENBLREL. 15, Created)
DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 15, Last annotation update)
DE Matrix extracellular phosphoglycoprotein precursor.
GN MEPE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Carnivora; Homniidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP TISSUE=BONE.
RC MEDLINE=20399567; PubMed=10945470;
RA Rowe P.S.N., De Zoysa P.A., Dong R., Wang H.R., White K.E.,
RT Econs M.J., Oudet C.L.;
RT "MEPE, a new gene expressed in bone marrow and tumors causing
osteomalacia.";
RL Genomics 67:54-68(2000) ✓
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=21309068; PubMed=11414762;
RA Argiro L., Desbarats M., Glorieux F.H., Ecarot B.;
RT "Mepe, the gene encoding a tumor-secreted protein in oncogenic
hypophosphatemic osteomalacia, is expressed in bone.";
RL Genomics 74:342-351(2001).
DR EMBL: AJ2276396; CAB97250.1; -;
DR EMBL: AF325916; AAK70343.1; -;
KW SIGNAL.
FT SIGNAL.
FT CHAIN 18 525
SQ SEQUENCE 525 AA: 58419 MW: 0977CA6E871CA9E5 CRC64:
Query Match 100.0%; Score 126; DB 4; Length 525;
Best Local Similarity 100.0%; Pred. No. 4,4e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 TDLQERGDNDISPFSGDGQPFKD 23
    |||||
Db 242 TDLQERGDNDISPFSGDGQPFKD 264
```

RESULT :
Q95KG5

ID	095K65	PRELIMINARY;	PRT;	500 AA.
AC	095K65			
DT	01-DEC-2001 (TrEMBLrel. 19, Created)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)			
DE	Hypothetical 55.6 kDa protein.			
OS	Macaque fascicularis (Crab eating macaque) (Cynomolgus monkey).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;			
OC	Cercopithecinae; Macaca.			
OX	NCBI_TaxID=9541;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=TEMPORAL LOBE RIGHT;			
RA	Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.			
RA	Suzuki Y., Sugano S., Hashimoto K.;			
RT	"Isolation of full-length cDNA clones from macaque brain cDNA			
RT	libraries.";			
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AB060891; BAB46894.1; -.			
KW	Hypothetical protein			
SQ	SEQUENCE 500 AA; 55577 MW; 9180265ADBEDC7BC CRC64;			

Query Match	97.68;	Score 123;	DB 6;	Length 500;
Best local similarity	0.78	100	10	10

Best Local Similarly 95.7%; Pred. No. 1.2e-10;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0

```

Qy      1  TDLQERGDNDISPFSGDGQPFKD  23
          |||||:|||||
Db      218 TDLQERGDNDISPFSGDGQPFKD  240

```

RESULT :
Q9GMI3

ID	ORGANISM	PRELIMINARY	PRG	555 AA.
DT	01-MAR-2001 (Tremblrel. 16, Created)			
DT	01-MAR-2001 (Tremblrel. 16, Last sequence update)			
DT	01-MAR-2001 (Tremblrel. 16, Last annotation update)			
DE	MEPE protein.			
OS	Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Macaca.			
OC	Cercopithecoidea; Macaca.			
OX	NCBI_TaxID=9541;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=BRAIN PARITIAL LOBE;			
RA	Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terai			
RA	Suzuki Y., Sugano S., Hashimoto K.;			
RT	"Isolation of full-length cDNA clones from macaque brain cDNA			
RT	libraries.";			
RL	Submitted (OCT-2000) to the EMBL/Genbank/DBDJ databases.			
DR	EMBL: AB050259; BAB17010.1, --			
SQ	SEQUENCE 555 AA; 61950 MW; BD4D47E88038A9E2 CRC64;			

Query Match	97.68;	Score 123;	DB 6;	Length 555;
Post Total Similarity	95.78;	Score 123;	DB 6;	Length 555;

Matches	22;	Conservative	1;	Mismatches	0;	Indels	0;	Gaps	0
---------	-----	--------------	----	------------	----	--------	----	------	---

```
Qy      1  TDLQERGDNDISPFSGDGQPFKD  23
          |||||:|||||
Db     273  TDLQERGDNDMSPFSGDGQPFKD  295
```

RESULT 4
Q9N076

AC 09N076; 2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Matrix extracellular phosphoglycoprotein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL: AB046056; BAB01638.1; -
SQ SEQUENCE 555 AA; 61979 MW; CCFE1A98ADA19EE4 CRC64;

Query Match	97.68;	Score 123;	DB 6;	Length 555;
PostgreSQL	97.68;	Score 123;	DB 6;	Length 555;

Best Local Similarity 95.7%; Pred. No. 1.4e-10;
Matches 22; Conservative 1; Mismatches 0; Indels 0; Gaps 0.

Qy	1	TDLQERGDNDISPFSGDGQPFKD	23
		:	
Db	273	TDLQERGDNDISPFSGDGQPFKD	295

RESULT 5
Q92411

AC	Q92411;	19, Created
DT	01-DEC-2001 (TREMBlurel.	19, last sec
DT	01-DEC-2001 (TREMBlurel.	19, last sec

DE	01-Jul-2002 (TrEMBLrel. 21, last annotation update)
DE	Matrix extracellular phosphoglycoprotein precursor.
OS	MEPE.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_Taxid:10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=C57BL/6J;
RX	MEDLINE=21309068; PubMed=11417472;
RA	Argito L., Desbarats M., Glorieux F.H., Ecartot B.,
RT	"Mepe, the gene encoding a tumor-secreted protein in oncogenic
RT	hypophosphatemic osteomalacia, is expressed in bone.";
RL	Genomics 74:342-351(2001).
DR	EMBL: AF314964; AAK70342.1; -.
DR	MCD; MGI:2137384; Mepe.
KW	Signal.
FT	SIGNAL
FT	CHAIN
SO	SEQUENCE
	1 18
	19 433
	433 AA; 45984 MW; 7CDB03CAACB41B0 CRC64;
	POTENTIAL.
	MATRIX EXTRACELLULAR PHOSPHOGLYCOPROTEIN

Query Match	64.38;	Score 81;	DB 11;	Length 433
-------------	--------	-----------	--------	------------

Matches	15;	Conservative	1;	Mismatches	4;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	----	--------	----	------	----

```
QY      2 DLQERGNDISPFGDGF 21
          || ||||: ||||| |
Db      171 DLVRGDNVPPFSGDGF 190
```

RESULT 6
Q9ES02

AC Q9ES02;


```

DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Osteoregulin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RX MEDLINE=20549633; PubMed=10967096;
RA Petersen D.N., Tkalcic G.T., Mansolf A.L., Rivera-Gonzalez R.,
RA Brown T.A.,
RT "Identification of Osteoblast/Osteocyte Factor 45 (OF45), a Bone-
RT Specific cDNA Encoding an RGD-containing Protein That Is Highly
RT Expressed in Osteoblasts and Osteocytes."
RL J. Biol. Chem. 275:36172-36180(2000).
DR EMBL: AF260922; AAG33366.1; -
SQ SEQUENCE 435 AA; 46515 MW; D587F82968A266CB CRC64;

Query Match 64.3%; Score 81; DB 11; Length 435;
Best Local Similarity 75.0%; Pred. No. 0.00037;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 DLQERGDNDISPFSGDCQPF 21
DB 165 DLVGRGNDVDPFSGDCQHF 184

RESULT 7
O52676 PRELIMINARY; PRT: 320 AA.
AC 032676;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE Porin.
GN PORC.
OS Rhodobacter capsulatus (Rhodospseudomonas capsulata).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Rhodobacter.
OX NCBI_TaxID=1061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=37B4;
RX MEDLINE=97149280; PubMed=8996088;
RA Trleschmann M.D., Pattus F., Tadros M.H.;
RT "Molecular characterization and organization of porin from Rhodobacter
RT capsulatus strain 37B4."
RL Gene 183:61-68(1996).
DR EMBL: U57653; AAB41301.1; -
DR HSSP: P31243; 2POR.
SQ SEQUENCE 320 AA; 33396 MW; B0E1745CB9D86F93 CRC64;

Query Match 41.3%; Score 52; DB 2; Length 320;
Best Local Similarity 58.8%; Pred. No. 7;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 TDLQERGDNDISPFSGD 17
DB 112 TDLQERGDNDISPFSGD 128

RESULT 8.
O9VYV0 PRELIMINARY; PRT: 793 AA.
AC 09VYV0;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE CG14085 protein.
GN CG14085.

```

```

OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Suton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Duntov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Foster A., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibbegam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclib J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Sliker E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Sylvestre R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003517; AAF49175.1; -
DR FlyBase: FBgn0036859; CG14085.
SQ SEQUENCE 793 AA; 88674 MW; FFC0E3DF3E537C41 CRC64;

Query Match 39.3%; Score 49.5; DB 5; Length 793;
Best Local Similarity 61.1%; Pred. No. 49;
Matches 11; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

OY 7 GDNDISPFSG-DCQPFKD 23
DB 693 GDNDISPFSGDCQPFKD 710

RESULT 9
O9C091 PRELIMINARY; PRT: 752 AA.
AC 09C091;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE KIAA1772 protein (Fragment).
GN KIAA1772.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;

```

(1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21082932; PubMed-11214970;
 RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIX.
 The complete sequences of 100 new cDNA clones from brain which code
 for large proteins in vitro.";
 RL DNA Res. 7:347-355(2000).
 DR EMBL: AB051559; BAB21863.1;
 DR InterPro: IPR001969; Asparticase_site.
 DR PROSITE: PS00141; ASP_PROTEASE; UNKNOWN_1.
 FT NON_TER
 SQ SEQUENCE 752 AA; 82489 MW; 7BEA604AA83D2E4 CRC64;

Query Match 38.9%; Score 49; DB 4; Length 752;
 Best Local Similarity 57.1%; Pred. No. 55;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 4 OERGNDISPFGD 17
 |||: ||| ||
 Db 630 ORGDSVTFPGD 643

RESULT 10

OBY183 PRELIMINARY; PRT: 277 AA.
 AC OBY183;
 DT 01-MAR-2002 (Tremblrel. 20, Created)
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
 DE Lysosyme M1 precursor (EC 3.2.1.17).
 GN BME10562.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID-29459;
 RP SEQUENCE FROM N.A.
 RC STRAIN-16M / ATCC 23456 / BIOTYPE 1;
 RX MEDLINE-20020109; PubMed-11756688;
 RA DelVecchio V.G., Kapatal V., Redkar R.J., Patra G., Mujer C., Los T.,
 RA Ivanova N., Anderson I., Bhattacharya A., Lykdis A., Resnik G.,
 RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Jettison J.J.,
 RA Haselkorn R., Kyrides N., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 Brucella melitensis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 DR EMBL: AE009498; AL51743.1;
 DR InterPro: IPR002053; CH_25.
 DR Pfam: PF01183; Glyco_hydro_25; 1.
 DR ProDom: PD004620; GH_25; 1.
 DR Hydrolase; Glycosidase; Complete proteome.
 SQ SEQUENCE 277 AA; 31766 MW; B8ICF4BA63E2B80 CRC64;

Query Match 38.1%; Score 48; DB 16; Length 277;
 Best Local Similarity 52.9%; Pred. No. 25;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 6 RGNDISPFGDQPFK 22
 ||| |||: ||| ||
 Db 252 RGDADINTFAGDSASWK 268

RESULT 11

O92432 PRELIMINARY; PRT: 385 AA.
 AC O92432;
 DT 01-NOV-1998 (Tremblrel. 08, Created)
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE LEF-3-ACMNPV orf67.
 GN LEF-3.

OS Bombyx mori nuclear polyhedrosis virus (BmNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID-10458;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-T3;
 RX MEDLINE-97329351; PubMed-9185864;
 RA Kamita S.G., Maeda S.;
 RT "Sequencing of the putative DNA helicase-encoding gene of the Bombyx
 mori nuclear polyhedrosis virus and fine-mapping of a region involved
 in host range expansion.";
 RL Gene 190:173-179(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-T3;
 RX MEDLINE-99281911; PubMed-10355780;
 RA Gomi S., Majima K., Maeda S.;
 RT "Sequence analysis of the genome of Bombyx mori
 nucleopolyhedrovirus.";
 RL J. Gen. Virol. 80:1323-1337(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-T3;
 RA Maeda S.;
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-T3;
 RA Gomi S.;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: L33180; AAC63740.1;
 SQ SEQUENCE 385 AA; 44875 MW; 1B2F1A12E88207BB CRC64;

Query Match 38.1%; Score 48; DB 12; Length 385;
 Best Local Similarity 56.2%; Pred. No. 36;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 8 DNDISPFGDQPFK 23
 ||| ||| ||| |||
 Db 61 DNKQEVYGGDSQSFKD 76

RESULT 12

O9F5V2 PRELIMINARY; PRT: 440 AA.
 AC O9F5V2;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE REB.
 GN REB.
 OS Bacteroides fragilis.
 OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Bacteroidaceae;
 OC Bacteroides.
 OX NCBI_TaxID-817;
 RP SEQUENCE FROM N.A.
 RC TRANSPOSON-CTN25285;
 RA Smith C.J., Bayley D.;
 RT "The putative IS1224 insertion sequence is actually part of a large
 genetic element with similarity to the Bacteroides conjugative
 transposons.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS A SIGMA-54 FACTOR INTERACTION ATP-BINDING
 DOMAIN.
 DR EMBL: AF303552; AAG17462.1;
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR002197; HTH_Fis.
 DR InterPro: IPR001789; Response_reg.
 DR InterPro: IPR002078; S1954_interact.
 DR Pfam: PF02954; HTH_8; 1.
 DR Pfam: PF00072; response_reg; 1.

DR Pfam: PF00158; Sigma54_activat; 1.
 DR PRINTS: PR01590; HTFIS.
 DR Prodom: PD000039; Response_reg; 1.
 DR SMART: SM00382; AAA; 1.
 DR SMART: SM00448; REC; 1.
 DR TIGRFAMs: TIGR01199; HTH_fis; 1.
 DR PROSITE: PS00676; SIGMA54_INTERACT; 2; 1.
 DR PROSITE: PS50045; SIGMA54_INTERACT; 4; 1.
 KM ATP-binding; DNA-binding; Phosphorylation; Sensory transduction;
 KW Transcription regulation.
 SQ SEQUENCE 440 AA; 49627 MW; 611F28F8A91E0799 CRC64;

Query Match 38.1%; Score 48; DB 2; Length 440;
 Best Local Similarity 52.0%; Pred. No. 42;
 Matches 13; Conservative 2; Mismatches 8; Indels 2; Gaps 1;

OY 1 TDQER--GNDISPFSGDGPFD 23
 DB 118 TLTKERNIGRMFVFSRDSAFKD 142

RESULT 13

OY 044497 PRELIMINARY; PRT; 484 AA.
 AC 044497; P94201;
 DT 01-NOV-1996 (TRENBLREL. 01, Created)
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
 DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)
 DE ALGJ protein.
 GN ALGJ.
 OS Azotobacter vinelandii.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Azotobacter.
 OX NCBI_TaxID=354;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-E.
 RA MEDLINE-97090405; PubMed-8936313;
 RA Rehm B.H.A.;
 RT *The Azotobacter vinelandii gene algJ encodes an outer-membrane
 RL protein presumably involved in export of alginate.*;
 RL Microbiology 142:873-880(1996).
 RN [2]
 RP SEQUENCE OF 1-128 FROM N.A.
 RC STRAIN-ATCC9046;
 RA Mejia-Ruiz H., Moreno S., Guzman J., Najera R., Soberon-Chavez G.,
 RA Espin G.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: X86533; CAA60245.1; -
 DR EMBL: X86533; CAA67371.1; -
 DR InterPro: IPR002106; AATRNA_Ligase1.
 DR PROSITE: PS00179; A_TRNA_Ligase1; UNKNOWN_1.
 SQ SEQUENCE 484 AA; 54611 MW; 0A4CCF18C928DBD3 CRC64;

Query Match 38.1%; Score 48; DB 2; Length 484;
 Best Local Similarity 47.4%; Pred. No. 47;
 Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 2 DUQERDNDISPFSGDGP 20
 DB 393 DDEDLGNGISPIEKDKP 411

RESULT 14

ID 09NR94 PRELIMINARY; PRT; 581 AA.
 AC 09NR94;
 DT 01-OCT-2000 (TRENBLREL. 15, Created)
 DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)
 DT 01-MAR-2001 (TRENBLREL. 16, Last annotation update)
 DE Hypothetical 64.2 kDa protein.
 BG:DS07295.4.
 GN Drosophila melanogaster (Fruit fly).

OC Eukaryota: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta:
 OC Pterygota: Neoptera: Endopterygota: Diptera: Brachycera: Muscomorpha:
 OC Ephydroidea: Drosophilidae: Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RA MEDLINE-99403001; PubMed-10471707;
 RA Ashburner M., Mista S., Roote J., Lewis S.E., Blazey R., Davis T.,
 RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
 RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
 RA Palazketo M., Reese M.G., Spreading A., Tsang G., Wan K., Whitelaw K.,
 RA Celniker S., Rubin G.M.;
 RT *An exploration of the sequence of a 2.9-Mb region of the genome of
 RT Drosophila melanogaster: the Adh region.*;
 RL Genetics 153:119-219(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RA Celniker S.E., Agbayani A., Arcalua T.T., Baxter E., Blazey R.G.,
 RA Butenhoif C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
 RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
 RA Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim E., Lee B.,
 RA Lewis S., Li P., Lomoten M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
 RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,
 RA Sechi H., Smir E., Svistkas R.R., Wan K.H., Weinburg T., Zhang R.,
 RA Zierman L.L., Rubin G.M.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE003412; AAF4927.1; -
 DR FLYBase: FBgn0028861; BG:DS07295.4.
 DR Hypothetical protein.
 SQ SEQUENCE 581 AA; 64222 MW; 44841FD16526701 CRC64;

Query Match 38.1%; Score 48; DB 5; Length 581;
 Best Local Similarity 47.1%; Pred. No. 58;
 Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 7 GNDISPFSGDGPFD 23
 DB 405 GANNLSPDPAHGHPVKD 421

RESULT 15
 OY 09KB7 PRELIMINARY; PRT; 444 AA.
 AC 09KB7;
 DT 01-OCT-2000 (TRENBLREL. 15, Created)
 DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
 DE Intracellular alkaline serine protease.
 DE APRX OR BH1930.
 OS Bacillus halodurans.
 OC Bacteria: Firmicutes: Bacillus/Clostridium group: Bacillales;
 OC Bacillaceae: Bacillus.
 OX NCBI_TaxID=86655;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C-125 / JCM 9153;
 RA MEDLINE-20512582; PubMed-11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT *Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis.*;
 RL Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL: AP001513; BAB05649.1; -
 DR HSSP: Q99405; IMPT.
 DR InterPro: IPR001128; Cytochrome_P450.
 DR InterPro: IPR000209; Peptidase_S8.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR PRINTS: PR00723; SUPRILISIN.
 DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.
 DR PROSITE: PS00136; SUPRILISIN_ASP; UNKNOWN_1.

Fri Jan 17 13:11:07 2003

us-09-641-034-47.rsp

Page 6

DR PROSITE: PS00137; SUBTILASE_HIS: 1.
DR PROSITE: PS00138; SUBTILASE_SER: 1.
KW Protease; Complete proteome.
SQ SEQUENCE 444 AA; 48916 MW; 3B05F81C53A21F2F CRC64;

Query Match	37.3%	Score 47	DB 16	Length 444
Best Local Similarity	47.1%	Pred. No. 61		
Matches	8	Conservative	4	Mismatches 5
				Indels 0
				Gaps 0

QY 2 DLQERGDNDISPFSGDG 18

Db 319 DTTDREDDVAFSSRG 335

Search completed: January 17, 2003, 13:18:38
Job time : 30 secs

Search completed: January 17, 2003, 13:18:38
Job time : 30 secs